



UNA

FENOMENOLOGÍA DE LA LIBERTAD

Un estudio sobre la libertad de la voluntad,
el concepto de responsabilidad moral
y la constitución del sujeto responsable

Javier Hernández Iglesias

Universidad Carlos III de Madrid

Director: Prof. Carlos Thiebaut Luis-André

ÍNDICE

AGRADECIMIENTOS.....	5
PRESENTACIÓN: OBJETIVOS Y PLAN DEL TRABAJO.....	8
INTRODUCCIÓN	15
1. UNA HISTORIA DEL CONCEPTO DE LIBRE ALBEDRÍO.....	15
El helenismo	18
La filosofía medieval.....	24
El Renacimiento	35
La filosofía moderna.....	40
La filosofía contemporánea.....	77
2. LA CUESTIÓN HOY	90
I- LA LIBERTAD DE LA VOLUNTAD	96
I-1. EL JARDÍN DE SENDEROS QUE SE BIFURCAN: EL REQUISITO DE LAS POSIBILIDADES ALTERNATIVAS.....	97
I.1.1. <i>El problema del determinismo</i>	101
I.1.2. <i>El problema del indeterminismo</i>	107
I.1.3. <i>Sobre la causalidad no determinística</i>	111
El indeterminismo sencillo	113
El indeterminismo causal	124
El agente causante.....	145
I.1.4. <i>La noción de “causa sui”, el “clinamen” y la tercera antinomia de la razón pura como destinos del libertarismo</i>	153
I-2. EL JARDÍN DE SENDEROS QUE SON FICTICIOS: LAS POSIBILIDADES ALTERNATIVAS DENTRO DE LOS LÍMITES DE LA MERA RAZÓN.....	168
I.2.1. <i>La crítica a las posibilidades alternativas</i>	169
I.2.2. <i>Qué hacemos cuando decimos “podría haber obrado de otro modo”: la forma cotidiana de ser determinista y contar con posibilidades alternativas</i>	178
I-3. EDIPO NO PODÍA SER SINO EDIPO... PERO EL ORÁCULO NO LO SABÍA: LA LIBRE VOLUNTAD Y EL AGENTE.....	199
I.3.1. <i>Necesidad y contingencia: determinación no es predeterminación</i>	200
I.3.2. <i>La personalidad y su carácter: unidad del agente y sujeto de la libertad</i>	218
RECAPITULACIÓN.....	229

II- LA RESPONSABILIDAD MORAL.....	231
II-1. EL DÍA A DÍA DE LA RESPONSABILIDAD MORAL.....	232
II.1.1. <i>La actitud participante y la actitud objetiva</i>	238
II.1.2. <i>La solución de Strawson al problema de la responsabilidad</i>	243
II.1.3. <i>La racionalidad de la responsabilidad moral</i>	245
II.1.4. <i>En torno al utilitarismo</i>	248
II-2. LA RESPONSABILIDAD: ¿NACE O SE HACE?.....	254
II.2.1. <i>Problemas con la estrategia naturalista</i>	256
II.2.2. <i>Problemas con la estrategia racionalista</i>	259
II.2.3. <i>¿Incompatibilidad de las dos estrategias?</i>	261
II-3. EL YO Y SU AUTOCONCIENCIA: FENOMENOLOGÍA DEL SUJETO RESPONSABLE.....	264
II.3.1. <i>La posibilidad de lo imposible</i>	265
II.3.2. <i>Un experimento fenomenológico</i>	268
II.3.3. <i>Limitaciones del naturalismo</i>	272
II-4. EL YO Y SU HETEROCONCIENCIA: VIVIR EN UN MUNDO PARTICIPANTE.....	278
II.4.1. <i>Una nueva versión de la estrategia racionalista</i>	280
II.4.2. <i>Mundo participante e incompatibilismo</i>	286
II.4.3. <i>Las razones de la vida participante</i>	291
RECAPITULACIÓN.....	296
III- EL MUNDO PARTICIPANTE.....	299
III-1. JU(Z)GAR CON LOS SENTIMIENTOS.....	301
III.1.1. <i>La actitud objetiva como posibilidad teórica</i>	303
III.1.2. <i>Las emociones reactivas</i>	306
La actitud objetiva y la investigación teleológica	307
Sobre las expectativas	309
La clase de las emociones reactivas	313
III.1.3. <i>Emociones reactivas morales y no morales</i>	317
III-2. ANATOMÍA DE LAS EMOCIONES RESPONSABLES.....	324
III.2.1. <i>Actitud participante y sentimientos reactivos</i>	325
III.2.2. <i>Hacia una taxonomía de las emociones reactivas</i>	327
III.2.3. <i>La indignación, la culpa y los límites de la moral</i>	335
III-3. EL ARTE DE EXCULPAR: ¿ QUÉ VALE Y QUÉ NO VALE?.....	342
III.3.1. <i>Emociones culturales y actitud reflexiva</i>	343
III.3.2. <i>Excusas y exenciones</i>	347
III-4. VIVIR EN UNA COMUNIDAD MORAL.....	356

<i>III.4.1. La personalidad y su virtud: las dos caras de la responsabilidad</i>	<i>357</i>
La responsabilidad personal	361
La responsabilidad areteica	362
<i>III.4.2. Los que viven fuera de la comunidad moral.....</i>	<i>366</i>
El cerebro del psicópata	367
Las emociones del psicópata	371
La gestación social del psicópata	377
<i>III.4.3. Aprender a participar</i>	<i>385</i>
RECAPITULACIÓN.....	391
CONCLUSIÓN.....	394
BIBLIOGRAFÍA	401

AGRADECIMIENTOS

Dado el punto de vista determinista que defiendo en estas páginas, no puedo sino admitir que esta obra es un efecto necesario de innumerables causas, destacando entre éstas todas las personas que, para bien o para mal, han formado parte de mi vida. Ahora, no todas las causas determinan por igual y asimismo no todas las personas que me rodean me han influido de igual modo a la hora de realizar este trabajo. Por ello no puedo dejar de mencionar aquí a algunas sin las cuales sé que esto no habría sido posible.

En primer lugar debo agradecerle cuanto me han enseñado a quienes para mí son los mejores filósofos que he conocido: Manuel Hernández Iglesias y Francisca Pérez Carreño. A ellos les debo, puede que a su pesar, el gusanillo de la filosofía y también cuanto de claro y riguroso pueda haber en mi discurso filosófico. Gracias a ellos comencé yo el camino que me ha traído hoy hasta este punto de mi carrera.

Indispensable fue también, para llegar hasta aquí, haber llegado antes a la Universidad Carlos III de Madrid, y a ella me atrajo la persona de Carlos Thiebaut, a quien tengo la suerte de poder admirar un poco más cada día como filósofo y como persona, y que ha sido y espero siga siendo un excelente tutor para mí además de un ejemplo de honestidad, tolerancia e infinita curiosidad intelectual (se me ocurren pocas virtudes más importantes en un filósofo).

Entre todas las cosas que nos influyen intelectualmente a lo largo de la vida siempre hay algunas que nos causan tal impresión que llegan a cambiar nuestro pensamiento de forma fundamental, esto me ha sucedido a mí en dos ocasiones. La primera en la Universidad Autónoma de Madrid: gracias a José Hierro Sánchez-Pescador (algo ayudado por Davidson y Putnam) aprendí que en el siglo XXI todavía tenía sentido dedicarse a la metafísica, y que era posible hacerlo con rigor y claridad. La segunda en la Universidad Carlos III: gracias a Carmen González Marín (algo ayudada por Wittgenstein) recuperé el mundo que me habían robado los escépticos y descubrí un tono filosófico, el pragmatista, productivo y optimista.

En general a la Universidad Autónoma le debo el descubrimiento de los principales problemas de la filosofía, y a la Universidad Carlos III muchas de sus soluciones o de sus inicios de solución. Todas las demás universidades deberían envidiar el fantástico grupo de filósofos con que cuenta el Área de Filosofía del Departamento de Humanidades y Comunicación de esta universidad, mi gratitud es inmensa para los profesores que lo componen, a saber (al margen de los ya citados Carmen y Carlos): Antonio Gómez Ramos, Antonio Valdecantos y Fernando Broncano. Vale más una hora de charla con cualquiera de ellos que muchas jornadas, simposios y congresos de filosofía, y lo mismo puede decirse de la charla con mis compañeros de doctorado. Muchas gracias a Rocío Orsi, Patricia Revuelta, Jaime Díez, Nuria Rivera, Clara Ramírez, Ana Fernández, David Conte, Tatiana Rincón, Alejandro Sahuí, Leticia Naranjo, Maya Siminovich, Noelia Alarcón, Álvaro Carvajal y muchos otros con los que he compartido sesudas conversaciones filosóficas y no tan sesudos momentos de ocio. Estoy convencido de que ambas cosas han sido igualmente productivas intelectualmente.

No puedo olvidarme aquí de mis compañeros y amigos de la Universidad Autónoma de Madrid. Fernando Comella tuvo que soportarme a su lado durante

toda la carrera (por lo que pude aprender de él tanto o más que de los profesores), para mí es un privilegio conocerle y haber compartido con él esos años de amistad y todos los que nos queden por compartir. No menos importante durante la carrera fue compartir también amistad e interés por la filosofía con María Gutiérrez Ariño, aproveché entonces y seguiré aprovechando siempre que pueda para disfrutar de su compañía, de su cariño y de su casi incombustible buen humor.

Por último he de citar a aquí a Elena Vacas, porque la filosofía es más una forma de vida que un corpus cerrado de conocimientos, así que con nadie puedo aprender más sobre filosofía que con quien comparto mi vida.

No hará falta que diga que, puesto que soy compatibilista, si bien todas estas personas han determinado el que este trabajo sea el que es, no por ello son responsables del mismo, sólo a mí corresponde esa carga. Y mi primera obligación como responsable de esta obra es de gratitud, por ello, muchas gracias a todos los que sabéis que estáis aquí, en este libro, y perdón si a algunos no he sido capaz de citarlos.

No querría dejar de señalar que la elaboración esta tesis doctoral ha sido posible gracias a que he disfrutado de una Beca de Formación de Profesorado del Ministerio de Educación, Cultura y Deportes, que me ha permitido dedicarme en exclusiva a la investigación durante mis cuatro años de doctorado.

A mis padres, que me dieron la mejor educación posible y que para mí han sido, son y serán siempre un ejemplo de inquietud intelectual y espíritu crítico, un sano antídoto contra el fanatismo y el dogmatismo.

PRESENTACIÓN: OBJETIVOS Y PLAN DEL TRABAJO

Me propongo tratar aquí el problema filosófico de la libertad de la voluntad, y aunque la búsqueda de una definición clara y completa de aquello que podría entenderse por libertad de la voluntad es precisamente el objeto de gran parte de este trabajo, sería dejar coja esta presentación no dar una definición cuando menos aproximada del tema que va a ser estudiado. Por ello diré que no se trata aquí de investigar acerca de lo que Schopenhauer denomina la “libertad física”, “que hace referencia únicamente al *poder*, es decir, a la ausencia, justamente, de impedimentos *físicos* de sus acciones”¹. Este concepto de libertad consistente en el *poder* para hacer lo que uno quiera es sin duda el más manejado en la vida cotidiana, y en filosofía a menudo nos referimos a él mediante las expresiones “libertad política” o “libertad de acción”. Esta disparidad de expresiones muestra que no se trata de un concepto unívoco, pero resolver la equivocidad de ese concepto de libertad no constituye el objeto de este trabajo; baste sencillamente con señalar que aquí no me voy a ocupar de la libertad como un *poder* al margen de trabas externas (bien objetos, bien personas, bien leyes, etc...). Voy a ocuparme, en cambio, de la “libertad moral”, por seguir con la clasificación de Schopenhauer, esto es, una libertad de *querer*. El nombre filosófico más corriente para este concepto es el de “libre albedrío”, aunque también se habla de “libertad individual” o de “libertad de elección”. Aquí casi siempre hablaré de *libertad de la voluntad*. Dicha libertad no hace referencia a la posibilidad de realizar físicamente nuestras acciones, sino a la posibilidad de que la voluntad que nos mueve a la acción sea la que es, y no otra, solamente en virtud de sí misma. Dicho de otro modo: voy a ocuparme de averiguar hasta qué punto es cierto, no que podamos hacer

¹ Schopenhauer, *Sobre la libertad de la voluntad*, 2000: 46.

lo que queremos, sino que *podamos querer lo que queremos*, esto es, que seamos realmente dueños de nuestra voluntad.

Ahora, a pesar del carácter eminentemente abstracto de la cuestión que nos ocupa es mi intención tratar el tema, a la Kant, principalmente desde el punto de vista de la razón práctica. Por ello, si bien en un principio abordaré la cuestión desde sus fundamentos metafísicos en unos términos parecidos a los que he usado más arriba, más adelante me interesará sobre todo desarrollar una teoría acerca de cómo hacer operativa una determinada definición de la libertad de la voluntad. Esto último será realizado mediante el estudio del concepto de responsabilidad moral y de su uso, así como de la constitución del sujeto como persona responsable de sus acciones, capaz de dar cuenta de las mismas y de ser alabada o censurada por ellas. Para ello dividiré el trabajo en tres partes, a las que añado una introducción histórica.

La finalidad de la citada introducción es permitirnos rememorar las principales teorías que sobre la libertad han sido y las raíces conceptuales de muchas de las teorías actuales. Se trata pues, por recoger la afortunada distinción de Robert Brandom en *Tales of the Mighty Dead*, de una historia *de re* y no *de dicto*, una historia en que se “especifica el contenido conceptual diciendo qué es aquello *acerca de* lo que se está hablando, en el sentido normativo del objeto cuyas propiedades y relaciones uno necesita investigar en orden a aseverar la *verdad* de la explicación en cuestión.”²

Tras esta genealogía del concepto, la primera parte del trabajo aborda la eterna contienda entre compatibilistas e incompatibilistas, esto es, entre aquellos que defienden que para la libertad conlleva determinación y los que opinan, por el contrario, que libertad y determinismo son incompatibles. También estudia qué se puede entender por “libertad de la voluntad”, si dicha libertad es o no compatible con el determinismo y en qué consiste lo que

² Brandom 2002: 101.

normalmente se considera la condición *sine qua non* de la acción libre, a saber: las posibilidades alternativas. Intentará pues descubrir qué queremos decir cuando usamos las expresiones “podría haber actuado de otra manera” o “podría haber hecho otra cosa”.

La segunda parte investiga los fundamentos de la responsabilidad, a saber, la cuestión de, dada una determinada definición de la libertad de la voluntad, en qué consiste ser responsables de nuestras acciones, qué es lo que dota de sentido práctico a nuestras atribuciones de responsabilidad y por qué es adecuado vivir en un mundo transido por dichas atribuciones.

Por fin, la tercera y última parte estudia bajo qué formas solemos atribuir responsabilidad. Se interesa pues por la cuestión más directamente práctica de en qué ocasiones es o no correcto imputar responsabilidad y de cuáles son las características que definen a un sujeto responsable.

Resumiendo, puede decirse que en la primera parte trataré de hacer una fundamentación de la metafísica de la moral, en la segunda dicha fundamentación empezará a entretenerse con los pormenores de una metafísica de la moral propiamente dicha, y sólo en la última y tercera parte se verá ésta plenamente desarrollada, dejándose ya de lado definitivamente cuestiones de fundamentación y centrándose más en el aspecto normativo de la libertad.

No quedará más que concluir resumiendo las consecuencias más importantes de la teoría de la responsabilidad que se haya ido construyendo.

Este es pues el plan del trabajo. En cuanto al método hay que precisar que, si bien el título de este estudio reza *Fenomenología de la Libertad*, sólo puede considerarse que haré un estudio fenomenológico en cierto sentido que no puedo dejar de concretar aquí: lo que voy a llevar a cabo es una fenomenología *more* pragmatista (o un trabajo pragmatista *more* fenomenológico).³ Se trata de

³ Probablemente puede decirse que en realidad pretendo seguir un camino fenomenológico similar al que propone Paul Ricoeur en *La semántica de la acción*. Sin duda existen diferencias notables

un ejercicio de, tras la estela del *Individuos* de Peter Strawson, “metafísica descriptiva”. No hablo de fenomenología porque vaya a seguir fielmente el sendero metodológico y temático de Husserl, Sartre o Merleau-Ponty, sino sencillamente porque aspiro a explicar *cuál es la libertad que se da de hecho*, cuál es el concepto de libertad que subyace a nuestro comportamiento cotidiano, o dicho de otro modo: en qué consiste el fenómeno de la libertad. Renuncio de antemano a demostrar que la que yo suministre sea la única explicación posible de dicho hecho, lo cual no quiere decir que no vaya a criticar ciertas concepciones tradicionales de la libertad que no considero admisibles porque necesitan, por ejemplo, del dualismo de sustancias o de la hipótesis de la *causa sui*. Pero entiendo que desde alguna perspectiva de corte menos materialista podría llegarse a la conclusión de que aquello a lo que aquí voy a llamar libertad, debido a su carácter marcadamente causalista, merecería otro nombre. Eso sí, quedaría en manos de esos filósofos dar con una construcción coherente del todo libertad/responsabilidad que pueda dar cuenta del mismo número de manifestaciones de la libertad que aquí pretendo abordar.

Naturalmente, el que renuncie a demostrar que la teoría de la libertad que aquí ofrezca es la única posible no implica que no vaya a dar argumentos en su favor, tan sólo pretende dar cuenta de un espíritu falibilista que considero tanto más indispensable cuanto más metafísico es el problema a tratar. De no tenerse presente el falibilismo se corre el riesgo de sufrir un vértigo esencialista que fácilmente podría hacernos caer en afirmaciones más tajantes de lo que los seres humanos, con nuestra impresionante pero limitada facultad de juzgar, nos podemos permitir. En cualquier caso, como digo, sí daré argumentos, básicamente dos aunque con múltiples ramificaciones: en la primera parte se tratará de un argumento fenomenológico (con algún sostén lógico), adecuado sin duda al carácter más metafísico de esa parte del trabajo;

entre lo que hace Ricoeur y lo que aquí pretendo hacer, no obstante ambos partimos de la constatación de que fenomenología y análisis son, o podrían ser, dos caras de la misma moneda.

en la segunda parte, que versará sobre cuestiones más prácticas, daré un argumento de corte pragmatista.⁴ No considero que ambas estrategias sean incompatibles, pero posiblemente tampoco son complementarias; es mi impresión que cada uno de esos argumentos por separado es suficiente para justificar la teoría de la libertad que aquí se defiende, sin embargo su unión puede resultar bastante más persuasiva pues hay demasiados filósofos que dudan de la eficacia de los argumentos fenomenológicos y otros tantos que dudan de la eficacia de los argumentos pragmatistas. Confío en que no haya tantos que duden de ambos a la vez, a menos que se trate de filósofos dispuestos a admitir cosas tales como el dualismo de sustancias, la noción de *causa sui* o algunos otros espectros metafísicos que, desatendiendo el consejo de Ockham, multiplican sin necesidad ni remedio el número de entes hasta redundar en sistemas donde hay demasiados entes y demasiado poco ser o un mundo de las ideas que se extiende por doquier y un mundo real inexistente. Como en el caso de quien postula la existencia de Dios, cae en los hombros de esos filósofos la carga de la prueba de explicar cómo funciona un mundo con dos sustancias o cómo puede algo ser *causa sui*. Aquí me voy a limitar a tratar de desarrollar una teoría de la libertad coherente con nuestra experiencia cotidiana del mundo, el que no sea coherente con otro tipo de experiencias más allá del mundo no me atañe. No soy tan ingenuo filosóficamente como para considerar que hay certezas que no pueden ser puestas en cuestión; no obstante toda labor filosófica o científica debe partir de algún supuesto y aquí quiero partir de unos supuestos muy inmediatos como pueden serlo el monismo materialista y la ley de causa y efecto. No descarto que el desarrollo de las ideas que aquí se exponen sea compatible con posiciones dualistas o con

⁴ En realidad iré mezclando consideraciones fenomenológicas y pragmatistas a lo largo de todo el trabajo; no obstante es cierto que la carga del argumento en la primera parte recaerá sobre todo en un experimento fenomenológico que pretende dar cuenta de la experiencia de las posibilidades alternativas, y en la segunda en una serie de consideraciones acerca de la práctica cotidiana del lenguaje a la hora de atribuir responsabilidad.

la existencia de Dios, por ejemplo, pero en cualquier caso el dualismo y el deísmo no son mi punto de partida ni mi problema, sino el de sus defensores.

Dicho esto, y para facilitar al lector su tarea, no quisiera dejar de avanzar cuál es el tipo de teoría de la responsabilidad al cual querría llegar. En tanto se trata de dar cuenta de la libertad como fenómeno, aspiro a desvelar y ordenar coherentemente los conceptos que están implícitos en nuestra vida cotidiana cuando hacemos cosas tales como atribuir responsabilidad, castigar o alabar las conductas de los demás o incluso intentar corregirlas para el futuro. Todo ello, intentaré mostrar, supone que en nuestra forma de vida ya está presupuesta la libertad, y por lo tanto mi labor no será tanto la de fundamentar ese supuesto como la de ir desgranando sus manifestaciones hasta dar con las ideas en que se basa y estudiar la coherencia de las mismas entre sí, siendo consciente de que, si una de ellas no es correcta, entonces se debería concluir que el supuesto de la responsabilidad no tiene fundamento racional. No puedo por menos que anticipar aquí cuáles son esas ideas a fin de que el lector sepa cuáles son las hipótesis que no querría dejar de corroborar a lo largo de este trabajo. Éstas se comprenderán mejor más adelante, a medida que vaya avanzando el estudio en su ejercicio de clarificación y distinción de los diferentes conceptos presentes en el supuesto de la libertad de la voluntad. De momento me conformaré con que dichas hipótesis figuren aquí del siguiente modo:

- 1) Las acciones humanas están determinadas, determinadas por los propios seres humanos.
- 2) Una o más personas pueden actuar de forma diferente en situaciones muy parecidas.
- 3) Es posible predecir comportamientos, sin embargo las acciones humanas poseen un buen grado de imprevisibilidad.

- 4) Tenemos expectativas y emitimos juicios de responsabilidad sobre las acciones ajenas y propias basándonos en dichas expectativas.
- 5) Tenemos un mapa de excepciones que nos permite suspender esos juicios de responsabilidad.

En efecto, desde mi punto de vista basta con la verdad de estas afirmaciones para poder contar con una teoría de la libertad y la responsabilidad moral coherente y operativa. Las tres primeras hipótesis se refieren sobre todo al problema metafísico de fondo de la libertad de la voluntad y serán tenidas en consideración en la primera parte de este trabajo. La cuarta y la quinta me permitirán elaborar una teoría de la responsabilidad moral en las partes segunda y tercera, que tratarán sobre todo de reordenar coherentemente y de matizar la teoría desarrollada en apartados anteriores, así como mostrar su operatividad para resolver problemas concretos de atribución de responsabilidad.

Dicho esto pasaré a introducir el problema de la libertad de la voluntad con más precisión mediante un somero estudio del desarrollo histórico del mismo.

INTRODUCCIÓN

1. Una historia del concepto de libre albedrío

En virtud tan solo del siguiente hecho, a saber, que una controversia se ha mantenido en pie mucho tiempo y aún queda por dirimir, podemos suponer alguna ambigüedad en la expresión y el que los interlocutores están asignando ideas distintas a los términos utilizados en la controversia. [...] Éste ha sido el caso de la muy discutida cuestión sobre la libertad y la necesidad (...). Reconozco que esta disputa ha sido tan tratada por todos y ha llevado a los filósofos a un laberinto tal de oscura sofistería, que no sería extraño que un lector sensato satisfaga su necesidad de tranquilidad hasta el punto de hacer oídos sordos a la propuesta de tratar esta cuestión, de la que no puede esperar ni enseñanza ni entretenimiento.⁵

Muchos filósofos encontrarán tal vez frívola la pretensión de hacer en unas pocas páginas una historia de la filosofía, por ello es necesario aclarar aquí los motivos que hay detrás de estas páginas históricas. En primer lugar la, por otra parte, muy sana costumbre de muchos filósofos de abordar directamente un problema sin rastrear las raíces del mismo, lo cual permite que la reflexión filosófica no se enfangue de por vida en una época o un autor, esa costumbre digo, favorece que a veces se descubran mares mediterráneos disfrazados de nuevos y seductores océanos, y esto no es sino otra forma de enfangar una cuestión. En segundo lugar, las genealogías de conceptos pueden ayudar a descubrir contingencia donde nos habíamos acostumbrado a ver necesidad, y con ello favorecer el abandono de algunas ideas que parecían imprescindibles y que luego resultan ser fruto de algunos intereses históricamente locales que consiguieron imponerse de forma generalizada, no en virtud de su universalismo, sino de razones más espúreas. Hay conceptos, y sobre todo *versiones* de conceptos, que son como las muelas del juicio, que

⁵ Hume, *Investigación sobre el conocimiento humano* (Sección 8: “De la libertad y la necesidad”, Parte I), 1981: 104-105.

estaban sujetas a necesidades adaptativas pero hace tiempo ya que perdieron su función y ahora, o bien ni tan siquiera crecen y no se las echa en absoluto de menos, o bien crecen y, salvo excepciones, duelen y molestan hasta que son retiradas.

Como dije más arriba no se trata de una historia *de dicto*, esto es, una historia hermenéuticamente potente capaz de presentarnos a los agentes históricos sin apenas interferencias interpretativas acercándonos así a las palabras mismas de dichos agentes. Por el contrario el “estilo *de re* de la historiografía intelectual requiere apostar los *hechos* a lo largo de las demandas del texto, a la hora de extraer consecuencias, aseverar datos, y así delinear su contenido conceptual. La responsabilidad de justificar esas hipótesis auxiliares descansa en el adscriptor, más bien que en aquel al que le son adscritos los compromisos cuyos contenidos están siendo especificados.”⁶ También puede decirse, a la Hegel, que el ejercicio histórico que ocupa las siguientes páginas se encuadra en la “historia reflexiva pragmática”. En el primer capítulo de la introducción especial a sus *Lecciones sobre la filosofía de la historia universal*, Hegel analiza las distintas maneras de considerar la historia y señala estas tres: la historia inmediata, la historia reflexiva y la historia filosófica. A su vez, dentro de la historia reflexiva Hegel distingue entre la general, la pragmática, la crítica y la especial. En la manera *pragmática* de considerar la historia “[l]a necesidad de un presente se manifiesta siempre al espíritu; y este presente tiénelo el espíritu en el intelecto. El nexo interior de los acontecimientos, el espíritu general de las relaciones es algo perdurable, algo nunca caduco, algo presente siempre. Los acontecimientos son distintos, pero lo universal e interno, el nexo, es siempre uno.”⁷ En el caso que nos ocupa ese nexo es la libertad de la voluntad, o cuando menos las dos principales versiones que de la misma ha habido: la libertarista y la compatibilista. Dichas lecturas de la libertad se han

⁶ Brandom 2002: 102-103.

⁷ Hegel, *Lecciones sobre la filosofía de la historia universal*, 1989: 157.

perpetuado desde el helenismo hasta nuestros días, y la historia que recoge este hecho como un continuo “anula el pasado y hace presente el acontecimiento. Las reflexiones pragmáticas, por abstractas que sean, resultan efectivamente algo presente e insuflan vida actual a las referencias del pasado.”⁸ Por ello, en esta introducción histórica discutiré con el pasado con la voz del presente, esperando devolver a los protagonistas de la genealogía del concepto de libertad al debate que acerca del mismo tiene lugar en nuestros días.

Dicho esto, tan sólo añadiré que apenas he consultado bibliografía secundaria para elaborar esta introducción. He intentado que los autores, en la medida de lo posible, hablen por sí mismos y así he evitado verme inmerso, sin quererlo, en luchas encarnizadas entre doxógrafos enemigos. Una vez hechas estas aclaraciones pasaré ya a relatar los avatares históricos de la libertad.

Ya Platón en su magna obra la *República* construye una fábula, el mito de Er, destinada a apuntar algo así como que cada uno es responsable de ser quien es y que la forma de vida que uno lleva es fruto de una elección personal y no de la voluntad de los dioses. En efecto, según nos relata Platón, Er fue testigo de cómo las almas de los mortales, antes de volver a la vida, se encontraban con las tres Moiras, y Láquesis, la moira que canta las cosas pasadas, les presentaba todos los modelos de vida imaginables: “Que el que resulte por sorteo el primero elija un modo de vida, al cual quedará necesariamente asociado. En cuanto a la excelencia, no tiene dueño, sino que cada uno tendrá mayor o menor parte de ella según la honore o la desprecie; la responsabilidad es del que elige, Dios está exento de culpa.”⁹

⁸ Ibidem.

⁹ Platón, *República*, Libro X (617e).

También en Aristóteles está presente el tema de la libre voluntad, al que el filósofo dedica gran parte del libro III de la *Ética Nicomáquea*, del Libro I de la *Gran Ética* y del libro II de la *Ética a Eudemo*, donde afirma que “(...) además de esto el hombre es el único animal que es, asimismo, principio de ciertas acciones; en efecto, no podemos decir de ningún otro animal que actúa”¹⁰ y también que “cada hombre es causa de las acciones voluntarias y conformes a su libre elección, pero no de las involuntarias; y todo lo que hace después de haberlo deliberadamente escogido, está claro que lo hace voluntariamente.”¹¹ Aristóteles es el primero en hablar de cosas que dependen de nosotros (*eph'hèmin*): “Pero si esto es evidente y no tenemos otros principios para referirnos que los que están en nosotros mismos, entonces las acciones cuyos principios están en nosotros dependerán también de nosotros y serán voluntarias.”¹² Y esta idea de cosas que dependen de nosotros en que se basaría la voluntad está más que cerca de nuestro moderno concepto de responsabilidad, en el que se ha focalizado el problema de la libertad de la voluntad en los últimos decenios. De hecho la teoría de Aristóteles acerca de los actos voluntarios y de cuando es o no lícito alabar o censurar a alguien en razón de la voluntariedad de sus actos, de si dependen de él o no, está en gran parte vigente y habrá ocasión de volver a ella más adelante. Sirva aquí como resumen de la misma este pasaje de la *Ética nicomáquea*: “Siendo involuntario lo que se hace por fuerza o por ignorancia, lo voluntario podría parecer que es aquello cuyo principio está en el mismo agente (...).”¹³

El helenismo

Ahora, aunque se pueda leer a Platón y a Aristóteles teniéndose presente el concepto de libertad, en rigor el concepto de libre albedrío es una

¹⁰ Aristóteles, *Op. Cit.*, Libro II, (1222b 18).

¹¹ Ibidem (1223a 17).

¹² Aristóteles, *Ética nicomáquea*, Libro III (1113b 19).

creación del helenismo, lo cual dista de ser casual pues parece lógico que el tema de la libertad individual tome cuerpo precisamente en un momento histórico en que la mirada de los filósofos se vuelve hacia el individuo, cuando la *polis* ya no puede ser un referente absoluto para la vida moral de los ciudadanos y éstos deben buscar en sí mismos el fundamento de la virtud y la felicidad. Como bien nos recuerda Hegel en su *Fenomenología del Espíritu*: “Esta libertad de la conciencia de sí, que emerge en su manifestación consciente por sí misma en el curso de la historia, ha sido, como es bien sabido, llamada *estoicismo*.”¹⁴ El concepto de libre albedrío no nace sino en el marco de la discusión entre epicúreos y estoicos,¹⁵ en el marco de un mundo moral que no se reduce ya a las murallas de la ciudad, sino que se extiende a toda la sociedad humana, a la naturaleza, al cosmos. De hecho, los problemas que entraña el uso del concepto de libre albedrío están directamente vinculados a las diferentes concepciones del mundo desde el punto de vista físico que sostienen los miembros de las citadas escuelas: a los epicúreos la doctrina de Demócrito les obliga a sostener una idea de libertad que encaje con el universo atómico, en el que podría caber cierta indeterminación; los estoicos, en cambio, han de desarrollar una idea de la libertad anclada en un mundo sometido a la necesidad implacable del destino. Pero no sólo la física influye de forma determinante en la moral de estoicos y epicúreos, también desde el punto de vista lógico existen divergencias entre dichas escuelas que están lejos de dejar sin consecuencias sus respectivas teorías del libre albedrío: los epicúreos no pueden admitir, como sí hacen los estoicos, la bivalencia del valor de verdad de las proposiciones, ya que esto les llevaría a admitir que toda predicción se refiere a un futuro inevitable, pues para que cualquier predicción pueda ser considerada verdadera o falsa antes de verse corroborada, ha de

¹³ Ibidem (1111a 23).

¹⁴ Hegel, *Phénoménologie de l'Esprit*, I, 1999: 169.

considerarse invariable el estado de cosas futuro que atañe a la predicción, de lo contrario una proposición verdadera podría, con el tiempo, convertirse en falsa y viceversa.¹⁶

En los términos actuales en que se trata la cuestión determinismo/libertad (como se verá más detenidamente en el siguiente apartado de esta introducción), se podría decir que la doctrina epicúrea es incompatibilista y la estoica compatibilista. Como ejemplo de la concepción incompatibilista o no-causal de la libertad se pueden recoger estas palabras de Epicuro en su *Carta a Meneceo*: “(...) [el hombre superior] se burla del destino, considerado por algunos señor absoluto de todas las cosas, afirmando que algunas suceden por necesidad, otras casualmente; otras, en fin, dependen de nosotros, porque se da cuenta de que la necesidad es irresponsable, el azar inestable, y, en cambio, nuestra voluntad es libre y, por ello, digna de merecer repulsa o alabanza”¹⁷; o también se pueden citar los más tardíos versos de Lucrecio: “(...) además de que nuestros movimientos / ni a tiempos ni a lugares se sujetan / determinadamente ; su principio / es nuestra voluntad (...)”.¹⁸ Como se ve los epicúreos niegan que la libertad pueda ser fruto de la necesidad. Esta idea es problemática si se atiende a la concepción fisicista del universo atómico que tienen los epicúreos: en principio todo movimiento a nivel macroscópico es fruto de la deriva infinita y *necesaria* de los átomos hacia espacios vacíos, pero la libertad puede, según Epicuro, convivir con el determinismo, sólo si *éste no implica necesidad* alguna. Esta es la razón de que

¹⁵ Aunque pronto la cuestión interesará a filósofos helenísticos de otras escuelas que se ocuparán de la misma con notable maestría (véase, por ejemplo, el caso de Alejandro de Afrodísia).

¹⁶ La discusión a este respecto es fruto del conocido problema que plantea Aristóteles en su *Sobre la interpretación* acerca de la atribución de valor de verdad a enunciados del tipo “mañana habrá una batalla naval”, problema del que también se ocupa, entre otros, Cicerón en *El hado*, eso sí, con una notable diferencia: Aristóteles cree estar resolviendo un problema lógico, desvinculado de problemas morales, no así Cicerón y los estoicos. A este respecto véase Anscombe 1956.

¹⁷ Epicuro, *Op. Cit.* [133]. Es cierto que en esta cita da la impresión de que la libertad según Epicuro no está ligada al azar, sin embargo esta impresión es fruto de que se trata de un texto de ética, dentro de su física Epicuro se ve obligado a hacer depender la libertad de un movimiento atómico de carácter impredecible por su aleatoriedad.

Epicuro, según Cicerón, “pida a sus átomos libertar el mundo, los separe de su camino [y le lleve a] suponer hechos sin causa (...) [y] admitir que de dos átomos llevados al vacío, uno sigue la línea recta, y el otro se separa de ella por sí mismo”¹⁹.

Pero como ya dije más arriba el problema al que se enfrentan epicúreos y estoicos no es únicamente físico, sino lógico-lingüístico. Los epicúreos deben criticar el principio semántico de bivalencia porque dicho principio entraña necesidad, y esa necesidad lógica incomoda a los epicúreos porque no da cabida a la idea de posibilidad, y para ellos la libertad radica en la *posibilidad* de hacer que algo sea o que no sea (de ahí que Epicuro no tenga nada contra el determinismo siempre que éste no implique necesidad). Así, la voluntad sería algo que podría contrarrestar la acción de las propiedades de peso y forma de los átomos haciendo que éstos lleven a cabo movimientos ajenos a los que sus propiedades intrínsecas les obligan a acometer: “Epicuro imaginó esta declinación [*clinamen*], porque temía que si la gravedad sola arrastraba a los átomos con movimiento natural y necesario, nada quedase libre en nosotros, moviéndose el ánimo según el impulso de los átomos”²⁰. Así quedaría resuelta la cuestión para los epicúreos en tanto habría cabida para posibles movimientos imprevisibles, alternativos a la deriva principal de átomos en el vacío y, por lo tanto, ajenos a necesidad alguna.²¹

Nada más lejos del concepto estoico de libertad que esta caracterización epicúrea de la misma: en primer lugar porque para los estoicos toda causa tiene un antecedente (véanse a este respecto las palabras de Crisipo que cita Cicerón en *El hado*: “(...) no puede realizarse ningún acontecimiento futuro sin que en el presente tenga causas por las cuales se realice alguna vez;

¹⁸ Lucrecio, *De la naturaleza de las cosas*, Libro II (260).

¹⁹ Cicerón, *La adivinación – El hado*, IX, 2002: 116.

²⁰ Ibidem: 117.

²¹ Más adelante, en I.I.4. habrá ocasión de volver a hablar de Epicuro y del *clinamen*, pues considero que algunas teorías libertaristas actuales no son sino una reedición del epicureísmo.

todo acontecimiento ha de tener causas con las que se encuentra necesariamente ligado”²²); en segundo lugar porque según las leyes de la lógica una proposición que es verdadera no puede no serlo (de nuevo Crisipo: “(...) todo lo que es verdadero de antemano se realiza fatalmente”²³); y en tercer lugar porque si interesa la libertad no es por la aparente aleatoriedad de la acción libre, sino, por el contrario, por la determinación que la rige. Así lo dispuso el fundador de la Stoa, Zenón de Citio, sentando las bases de una concepción compatibilista o causal de la libertad de la voluntad, como puede verse en estas palabras: “Es, en efecto, el destino la causa que produce concatenadamente los entes o la razón según la cual el mundo es gobernado”²⁴ y también “las causas de los acontecimientos en cierta medida [dependen] de nosotros y en cierta medida, no”²⁵. Y algo similar se puede ver en posteriores defensores del estoicismo²⁶ como Séneca, que señala en *Sobre la providencia*: “No se me obliga a nada, no sufro nada en contra de mi voluntad, no soy esclavo de Dios, sino que doy mi asentimiento, tanto más cuanto que sé que todo se desenvuelve de acuerdo a una ley fija, dictada para la eternidad. Los hados nos conducen (...). Una causa depende de otra. Una concatenación de hechos condiciona los acontecimientos privados y públicos.”²⁷

En su *Lógica del sentido* Gilles Deleuze ofrece una buena síntesis de todo lo que acabo de decir: “(...) la libertad es salvada de dos formas complementarias: una vez en la interioridad del destino como unión de las causas, otra vez en la exterioridad de los acontecimientos como vínculo de los

²² Ibidem: 118.

²³ Ibidem.

²⁴ Los estoicos antiguos, *Fragmentos*, (Zenón de Citio 281): [S.V.F. I 175].

²⁵ Ibidem: [S.V.F. I 177].

²⁶ Hay que señalar por no faltar al rigor que Crisipo, aún siguiendo el camino compatibilista marcado por Zenón, defiende una variante del determinismo más moderada gracias a una distinción entre tipos de causa y necesidad en la que no puedo detenerme aquí.

²⁷ Séneca, *Op. Cit.* 6-7, en 1986: 25.

efectos. Es por ello que los estoicos pueden oponer destino y necesidad.²⁸ Los epicúreos utilizan otro análisis de la causalidad, que fundamenta también la libertad: conservan la homogeneidad de la causa y del efecto, pero recortan la causa según unas series atómicas cuya independencia las unas con respecto a las otras está garantizada por el *clinamen* —no ya destino sin necesidad, sino causalidad sin destino. En ambos casos se comienza disociando la relación causal, en lugar de distinguir tipos de causalidad como hacía Aristóteles o como hará Kant.”²⁹

Hay que decir que la idea que más comúnmente se tiene hoy en día de la libertad no responde exactamente a ninguna de estas dos teorías (aunque se acerca a la epicúrea), sino a un tercer patrón que es consecuencia directa del cristianismo aunque más adelante haya sido secularizado: se trata de la idea del libre arbitrio que defiende Agustín de Hipona. Aunque en justicia debe señalarse aquí que dicha teoría ya fue en parte anticipada por Carnéades, que hizo un esfuerzo por zafar a la doctrina epicúrea de la libertad, de las críticas estoicas y defendió, según relata Cicerón, que “no debe buscarse causa externa al movimiento voluntario del ánimo, porque la naturaleza del movimiento voluntario lleva consigo que esté en nuestro poder y dependa de nosotros: no carece por consiguiente de causa, pero la causa está en su naturaleza misma.”³⁰ No obstante, hay diferencias notables entre la hipótesis de Carnéades y la posición de Agustín: la más importante de ellas es que el primero es materialista y no así el segundo... de hecho Cicerón acaba demostrando en su *El hado* que la hipótesis de Carnéades, por otra parte interesado en refutar a Crisipo más que en defender a Epicuro, termina más bien por conducir a una variación de la doctrina estoica de la libertad, compatible con el determinismo y la necesidad pero no con la idea de destino, que a una vertiente novedosa del

²⁸ Esto es sólo cierto en el caso de Crisipo, que maneja una noción débil de determinismo, y es en él en quien está pensando Deleuze, que sobre todo se basa en su texto en el *El hado* de Cicerón.

²⁹ Deleuze 1969 (2002): 15-16.

epicureísmo. Esto se debe sin duda a la dificultad que entraña casar la idea de *causa sui* con un universo fisicista, dificultad que no está presente en Agustín dado su dualismo platónico.³¹ Seguramente no es casual que Carnéades defienda esas doctrinas siendo el máximo representante de la Academia platónica en sus años de camino por la disciplina escéptica, disciplina que él contribuyó a desarrollar llevando los límites de la duda más allá de lo que lo habían hecho su inmediato antecesor Arcesilao y el propio Pirrón.

La filosofía medieval

Probablemente el giro más importante que se ha producido en la historia de la filosofía en torno al concepto de libertad es el que da Agustín, hasta el punto de que podría llegar a defenderse que es él quien convierte el tema de la libertad de la voluntad en uno de los problemas filosóficos más importantes. De hecho, cabe decir que la discusión entre epicúreos y estoicos es más que nada una discusión de ciencia de la moral, pues la divergencia existente entre ambas escuelas radica en una controversia referente a la composición y ordenación física del mundo, que a su vez tiene consecuencias para la libertad. Agustín de Hipona, sin embargo, no se preocupa tanto por el hecho de la libertad, por la realidad de la acción, como por sus fundamentos. En fin, Agustín convierte un problema físico en una cuestión metafísica, transforma un problema centrado en los actos en una cuestión relativa a las potencias... y este giro supondrá una redefinición de la cuestión de tal calibre que marcará de forma definitiva el posterior desarrollo de la misma.

La especial atención que le presta Agustín al concepto de libre albedrío se debe, además de a la importante presencia en la religión cristiana de la idea de culpa (cuya influencia ya se había dejado sentir en el último estoicismo), a la

³⁰ Cicerón, *La adivinación – El hado*, XI, 2002: 118.

mayor dificultad conceptual con que tropieza el pensamiento cristiano, a saber: el problema del mal. Se pregunta Agustín: “(...) si el pecado procede de las almas que Dios creó, y las almas vienen de Dios, ¿cómo no referir a Dios el pecado, siendo tan estrecha la relación entre Dios y el alma pecadora?”³² Puede decirse, sin que con ello se corra riesgo alguno de parecer extravagante, que la solución a este problema es una nueva concepción de la libertad según la cual ésta no puede considerarse ni fruto del azar ni producto de la necesidad; la verdadera libertad, en cambio, se daría en un acto causal, sí, pero efecto de una causa incausada. Dicho de otro modo: acto libre sería aquel acto realizado por una causa fruto única y exclusivamente de la voluntad. “Pero ¿cuál puede ser la causa de la voluntad anterior a la misma voluntad? O esta causa es la misma voluntad, y entonces en ella tenemos la raíz que buscamos, o no es la voluntad y en este caso la voluntad no peca. Así pues, o la voluntad es la primera causa del pecado o la causa primera del pecado está sin pecado (...).”³³ Esto ciertamente puede recordar a la doctrina epicúrea, sin embargo (como ya señalé más arriba) existe una diferencia notable entre la postura de los epicúreos y la de Agustín: el materialismo. En efecto, dice Agustín refiriéndose al movimiento de la voluntad: “(...) por lo mismo que no dudamos de que el movimiento del alma hacia el pecado sea culpable, por lo mismo debemos negar que sea natural (...).”³⁴ Epicuro se ve obligado a admitir la presencia de unos átomos impredecibles, Agustín no. Para Agustín la voluntad, en tanto se asienta en un alma que posee algunas de las virtudes del alma pura del creador, puede, en cierto sentido *ex nihilo*, generar voliciones: “(...) de dónde le viene a la voluntad el movimiento por el que se aparta del

³¹ Probablemente algunos incompatibilistas encontrarán esta afirmación discutible, mal intencionada o puede que incluso ofensiva; mostrar que no es gratuita será uno de los objetivos de la primera parte de este trabajo.

³² Agustín, *Del libre albedrío*, Libro I, Cap. II. 5, en 1971: 217.

³³ Ibidem, Libro III, Cap. XVII. 49: 401.

³⁴ Ibidem, Libro III, Cap. I. 2: 341-342. Léase este extracto teniendo en cuenta la distinción existente en la física aristotélica entre movimientos naturales y violentos.

bien inmutable y se une a lo mudable [...] ¿podemos acaso decir que es Dios el autor del pecado? No, y en este caso no procedería de Dios este movimiento. ¿De dónde pues procede? / Si al hacerme esta pregunta te respondiera que no lo sé, quizá se apoderara de ti una mayor tristeza, y, no obstante, te diría la verdad, porque nadie puede saber *lo que no es nada*, lo que no tiene ser.”³⁵ Estaríamos pues en presencia de una suerte de voluntad creadora, propia de un tipo de incompatibilismo que hoy recibe normalmente el nombre de *libertarismo*. La asunción de esta posición por parte de Agustín, más adelante convertida en doctrina de la Iglesia, planteará graves problemas a los filósofos medievales y a la mayor parte de los filósofos modernos, porque se trata de una concepción de la libertad que la presenta como incompatible con el determinismo, pero que defiende que sin embargo ha de ser compatible con la omnisciencia y la voluntad divinas, esto es, con una predeterminación divina que por definición implica necesidad pero que deberá, por razones morales, ser conjugable con una libertad de las acciones humanas ajena a la necesidad.³⁶ Por ello Agustín postula un tipo diferente de principio del movimiento (del que ya se han visto algunas características), que será asumido como correcto por la práctica totalidad de los filósofos medievales y por una mayoría de filósofos modernos. Resulta difícil no ver en el siguiente texto de Agustín una primera formulación no problemática de la tercera antinomia kantiana:

[S]ea propio de la voluntad aquel movimiento por el que *ella* aparta del Creador su voluntad para entregarla a las criaturas (...). Por lo demás, es semejante al que hace caer la piedra a tierra en que, así como éste es propio de la piedra, así aquel es propio del alma; pero se diferencia en que la piedra no tiene en su mano el impedir el movimiento por el

³⁵ Agustín, *Del libre albedrío*, Libro II, Cap. XX. 54, en 1971: 336. La cursiva es mía.

³⁶ Quede claro que este es un problema de los libertaristas cristianos, no es equiparable a los problemas con que se enfrentan libertaristas actuales laicos (posiblemente muchos de ellos se negarían a considerar a Agustín como uno de los suyos).

que es arrastrada hacia abajo, mientras que el alma, como no quiera, nadie la obliga a moverse (...); y he aquí por qué el movimiento de la piedra es natural y el del alma voluntario.³⁷

Como se puede ver, Agustín, al igual que Kant, distingue entre dos tipos de causalidad: una natural y otra por libertad. No obstante existe una diferencia básica en la teoría de ambos pensadores: en tanto la principal preocupación de Agustín es el problema del mal, sólo le es necesario plantear la libertad humana como posibilidad. El libre albedrío es un don que les otorga a los seres humanos la condición de seres morales, una potencialidad dada por Dios para que los seres humanos se hagan acreedores de una recompensa final por su buena vida, pero en principio todo ser humano tiende al bien (tal es la naturaleza que Dios dispone para cada ser humano), si bien puede actualizar su potencial libre albedrío para tender al mal (y entonces, como se acaba de ver, su movimiento deja de ser natural). Si la voluntad humana es capaz de tender al bien e incluso de hacer el bien, después del pecado original, es por la gracia divina que, en tanto que actualización por intervención divina de la potencial bondad del uso de la libertad, es lo que hace que esa tendencia al bien deba considerarse movimiento natural de la voluntad. No es natural porque sea necesario, sino porque Dios, por obra de la gracia, lo hace efectivamente posible (de no ser por la gracia divina obrar bien tras el pecado original sería posible tan sólo lógicamente). “Todas las cosas deben, pues, en primer lugar todo el ser que tienen, en cuanto son naturalezas. En segundo lugar deben toda su posibilidad de ser mejores, si quieren serlo; deben todas las gracias que han recibido para querer serlo, y deben, en fin, todo cuanto deben llegar a ser.”³⁸ El dios de Agustín no es sólo un primer motor, es también principio rector por obra de la gracia, y por ello Dios es la semilla y el garante del bien, y

³⁷ Agustín, *Del libre albedrío*, Libro III, Cap. I. 2, en 1971: 342. La cursiva es mía: nótese que “ella” sustituye a “la voluntad”, lo que está diciendo Agustín es que la voluntad mueve a la voluntad.

³⁸ Ibidem, Libro III, Cap. XVI. 46: 397.

de él *sólo* emana el bien porque, contrariamente a lo defendido por Fausto y los maniqueos, el mal, y en esto es en lo se hace más patente la herencia platónica y neoplatónica de Agustín, no es sino mera privación, ausencia de gracia. En ningún caso el mal es una sustancia, pues entonces habría sido creada por Dios, y difícilmente podría ser cualquier otra cosa: el mal, dada la omnipotencia divina, sólo puede ser un no-ser. El dios de Agustín por una parte otorga al ser humano el poder de decidir, lo hace moral por obra del libre albedrío; por otra parte le ayuda, le da algo más que una posibilidad meramente especulativa para usar bien su libre arbitrio, le concede la gracia. Sostener lo contrario es caer en la herejía de Pelagio según la cual, para salvarse, al hombre le basta con la buena voluntad que lo impulsa a hacer el bien. Pero esto no se corresponde con la doctrina de Agustín que renuncia a la tradición griega, de la que es heredero Pelagio, que vincula la libertad a la autarquía, en favor del concepto cristiano de gracia como aquello que permite a los hombres hacer el bien. Por ello en Agustín, aunque el libre albedrío es inherente a la condición humana, no se dice de la misma manera cuando los hombres tienden hacia el bien que cuando tienden hacia el mal, y esto se aleja mucho, como decía más arriba, de lo que defenderá Kant que, aunque coincida con Agustín en que hay que distinguir entre causalidad y voluntad, cree que la libertad se dice igual para el bien que para el mal.

Esta senda abierta por Agustín será seguida por la mayor parte de los filósofos medievales aunque no por todos. Así Boecio, que por ello podría ser considerado un estoico tardío, prescinde de esa libertad absoluta de la voluntad; para él es obvio que la providencia divina rige el mundo y también las vidas de todos y cada uno de nosotros: “[El Destino] dirige también los actos y la suerte de los hombres valiéndose de una indisoluble cadena de causas que, por fuerza, han de ser inmutables, del mismo modo que la

Providencia, donde tiene su origen, es inmutable.”³⁹ Boecio defiende que para que el ser humano sea considerado responsable, es suficiente con una libertad basada en la razón, ya que “aquel que por su misma naturaleza puede servirse de la razón tiene la facultad de poder juzgar [...] así, quien está dotado de razón tendrá también la libertad para querer o no querer”⁴⁰, y basta también con una libertad anclada en la causalidad del mundo, pues según Boecio “si un acontecimiento sucede sin causa, lógicamente es como si saliera de la nada. Y si esto es imposible, también lo será la existencia del azar (...).”⁴¹ Nada más alejado de la solución agustiniana que, como se ha visto, en algún pasaje pone en la nada el origen de la voluntad.

Pero el camino de Boecio será transitado de forma minoritaria (salvo en lo que respecta a su concepción intelectualista de la voluntad, de factura aristotélica). Los pensadores cristianos más importantes, aun marcando ciertas diferencias y especialmente entre los siglos IX y XII, seguirán a grandes rasgos las ideas de Agustín.

Anselmo de Aosta ahonda en la idea de esa voluntad cuya raíz es ella misma, en esa causalidad que no entraña necesidad, hablando de acciones espontáneas: “Libremente pecó el ángel y el hombre, porque pecó libremente aquel que era tan libre que nada podía forzarle a pecar (...), sin presión ni necesidad de ninguna clase, pecó espontáneamente. Sin embargo ambos pecaron por propia determinación, que era libre, es decir, por el *poder que les hacía capaces* de no pecar y de no estar bajo la esclavitud del pecado; sino por el poder que tenían de pecar, que no ayudaba a la libertad de no pecar ni sujetaba a la esclavitud del pecado.”⁴² Aquí se confirma lo que fue dicho más arriba, a saber, que Agustín abre la cuestión al territorio de las potencias, pues Anselmo

³⁹ Boecio, *La consolación de la filosofía*, Libro IV, 6, 1999: 150-151.

⁴⁰ Ibidem, Libro V, 2: 166.

⁴¹ Ibidem: 164.

en este texto está sobre todo esforzándose por distinguir entre *ser capaz* de pecar o no pecar, y *tener poder de ser capaz* de pecar o no pecar, o lo que es casi lo mismo: entre algo así como una libertad relativa (a las posibilidades reales) y una libertad absoluta (atendiendo a lo lógicamente posible). Y en la libertad absoluta es donde se centrará la discusión una y otra vez a lo largo de la historia del pensamiento, no obstante el esfuerzo de muchos filósofos por intentar sacar al concepto de libertad de ese atolladero en que lo metió Agustín, un atolladero que lleva a Anselmo a defender lo siguiente: “(...) si decimos que alguien, queriendo mentir para no perder la vida o la salud, se ve obligado a traicionar la verdad por temor a la muerte o de los tormentos, no decimos la verdad, porque no está más obligado a querer la vida que la verdad; pero como le está prohibido, por una fuerza extraña, conservar una y otra a la vez, escoge él mismo lo que ama más, espontáneamente y no contra su voluntad (...).”⁴³ Esto contradice plenamente la idea aristotélica de que no actuamos según nuestra voluntad cuando actuamos por ignorancia o por fuerza mayor, y también contradice nuestras intuiciones más inmediatas acerca de la responsabilidad o, como dice Aristóteles, acerca de lo que depende de nosotros; sin embargo sigue al pie de la letra la siguiente idea de Agustín: “Tampoco puedes ser violentado a hacer algo contra tu voluntad, porque tu voluntad no es menor que tu poder.”⁴⁴ Se convertirá pues en un problema clave de la filosofía medieval, dada la progresiva aristotelización del pensamiento cristiano a partir del siglo XII, el compatibilizar una idea absoluta de libertad (que inevitablemente conduce a sostener, como defiende Anselmo, que los seres humanos *siempre* que eligen, eligen libremente) con la ética aristotélica.

⁴² Anselmo de Aosta, *Del libre albedrío*, II, en 1952: 555. La cursiva es mía.

⁴³ Ibidem, IX: 579.

⁴⁴ Agustín, *Confesiones*, Libro VII, 4, 1990: 174.

Uno de los primeros pensadores en encontrarse con ese problema es Pedro Abelardo. En su *Ética*, aunque no llega a ahondar en el problema del libre albedrío, ciñéndose en lo fundamental a la teoría agustiniana, sí trata extensamente el tema de la voluntariedad de la acción humana, reavivando con ello la concepción aristotélica de la acción moral. Así, al mismo tiempo que sostiene que hay “muchas cosas que nos impiden actuar” pero que “la voluntad y el consentimiento siempre los tenemos a nuestra disposición”⁴⁵, Abelardo también sostiene que “no pecamos, por tanto, aunque deseemos o hagamos lo que no es lícito. A menudo estas cosas suceden sin que haya pecado alguno”.⁴⁶ Aunque con una dosis de internismo ajena a Aristóteles (pues la noción de pecado conduce irremisiblemente a tratar la buena o mala intención de las acciones), Abelardo coincide fundamentalmente con el filósofo griego en señalar que “a nadie se le oculta (...) con cuánta frecuencia hacemos cosas que no deben hacerse, sea por coacción, sea por ignorancia”⁴⁷ y que en esos casos no debe atribuirse responsabilidad moral.

Algo similar se puede descubrir en el pensamiento de Tomás de Aquino, si bien éste, a diferencia de Abelardo, sí trata con sistematicidad la parte metafísica del problema de la libertad. Así, en su *Suma Teológica* dice: “Y las acciones particulares son cosas contingentes, de modo que el juicio de la razón puede optar entre cosas diversas y no está determinado a una sola.”⁴⁸ Tomás lleva a cabo una intelectualización de la acción (acorde con la caracterización aristotélica de la deliberación moral) según la cual la libertad depende de la facultad de juzgar racionalmente (como también defendían Boecio y Pedro Abelardo, y como desarrollarán más extensamente Descartes y sus seguidores modernos). No obstante, el pasaje citado muestra que, en su

⁴⁵ Pedro Abelardo, *Ética o Conócete a ti mismo*, (Cap. III), 1990: 22.

⁴⁶ Ibidem: 21.

⁴⁷ Ibidem.

caracterización metafísica del libre albedrío, a grandes rasgos Tomás sigue a Agustín: no hay ninguna acción que nos esté dada *a priori* por la necesidad de la providencia divina, sino que “las acciones particulares son cosas contingentes” y “la voluntad como agente mueve todas las potencias del alma a sus actos”.⁴⁹ Tomás no llega a los límites de intelectualización de la acción a los que llega Boecio según el cual, como se vio más arriba, la capacidad racional de *distinguir* entre el bien y el mal y de *elegir* hacer el bien o el mal son lo mismo, lo cual es incompatible con la libertad absoluta agustiniana. Sin embargo Tomás, al igual que Pedro Abelardo, sí se aparta de Agustín desde el punto de vista de la articulación normativa de la responsabilidad que entraña la libertad (gracias a que Tomás distingue entre necesidad hipotética y necesidad absoluta, como habrá ocasión de ver más adelante, en *I.3.1.*). Así, Tomás sostiene que “(...) necesidad de coacción es la que repugna a la voluntad; pues llamamos violento a lo que es contra la inclinación de la cosa. (...) Así se llama a algo voluntario porque es según la inclinación de la voluntad. Por consiguiente, así como es imposible que una cosa sea al mismo tiempo violenta y natural, así es imposible que una cosa sea simplemente coaccionada o violenta y voluntaria.”⁵⁰ Y esto se parece mucho a la idea de acto voluntario que tenía Aristóteles, que rescata Pedro Abelardo y que, como dije más arriba, no es muy difícil defender aún hoy en día. Esta conciliación de la libertad absoluta agustiniana con una caracterización normativa relativa a las condiciones de la acción, es posible gracias a una sutil distinción entre dos formas de voluntad: la voluntad como fin (relativa y dependiente del entendimiento) y la voluntad como agente (absoluta y dependiente tan sólo de sí misma). En palabras de Tomás de Aquino: “(...) como fin; así se dice que el fin mueve al agente; y de esta manera el entendimiento mueve a la voluntad,

⁴⁸ Tomás De Aquino, *Summa teológica*, I p., L XXXIII, 1, en 1942: 247.

⁴⁹ *Ibidem*, 4: 245.

⁵⁰ *Ibidem*, 1: 238.

porque el bien entendido es el objeto de la voluntad y la mueve como fin; de otra manera, como agente (...) la voluntad mueve al entendimiento (...).”⁵¹ Naturalmente lo que pretende Tomás es adaptar la libertad agustiniana a la idea de ley natural y a la necesidad del fin último de la acción humana según Aristóteles, esto es, la felicidad. Así, los seres humanos actuarán conforme a su entendimiento a la hora de obtener los medios adecuados para procurarse la felicidad, pero actuarán libremente en tanto es su querer (libre, indeterminado) la felicidad lo que los mueve a pensar en qué medios son o no apropiados para alcanzarla.

Habría pues, por lo visto hasta aquí, dos grandes corrientes en la Edad Media en lo que respecta al tema del libre albedrío, pero estas corrientes no diferirían en lo esencial, esto es, en una concepción libertarista de la voluntad de corte agustiniano, sino en la ulterior articulación normativa de esa teoría. Así, habría un primer grupo de filósofos, más cercanos a Agustín, según los cuales difícilmente podría tenerse en cuenta eximentes de las malas acciones llevadas a cabo libremente; y habría también un segundo grupo de filósofos que aristotelizan a Agustín y elaboran sistemas morales menos rígidos a la hora de atribuir culpa. Como ya dije, parece escapar Boecio a este panorama, pero su influencia será escasa, considerado más como una variante cristiana de estoico que como un pensador cristiano por sí mismo con tanta autoridad como Agustín, autoridad que irá más allá de la Edad Media, en gran medida gracias al giro que da Duns Escoto al paradigma agustiniano.

La obra de Duns Escoto es indispensable porque viene a resumir los logros del pensamiento cristiano en lo que respecta al problema del libre albedrío, y deja la cuestión planteada en los términos en que se discutirá en la

⁵¹ Ibidem, 4: 244.

filosofía moderna: tanto Descartes, como Hobbes, Hume o Kant están discutiendo con Duns Escoto sin hacerlo explícitamente, y por ende indirectamente con Agustín. El Doctor Sutil será precursor de la doctrina de Descartes, que acabará distinguiendo entre una libertad genuina, sólo propia de Dios y que se confundiría con la idea de libertad de Agustín, y una libertad subsidiaria del ser humano, que sólo se da en el acto de pensamiento y que ya no tiene que ver con la de Agustín (haciéndose así eco Escoto de las posiciones más intelectualistas de Pedro Abelardo y de Tomás de Aquino). Lo que ha desaparecido es esa voluntad divina anclada en la necesidad de su propia esencia. En Duns Escoto la libertad divina es voluntad pura, y lo único que hace que Dios quiera algo es que tal es su voluntad: Dios únicamente está sometido a las leyes de la lógica, no al bien, porque el bien consiste precisamente en aquello que es fruto del acto puro de volición que es la voluntad divina. El bien ya no tiene ser por sí mismo, bien es aquello que Dios quiere que sea bien, ¿por qué? Porque lo quiere. Todo acto de Dios es un acto puro de voluntad, fruto de su *indiferencia*. Si la libertad es incompatible con la necesidad, entonces la voluntad de Dios no puede ser determinada ni aún por él mismo, pues algo determinado por Dios es algo que consta del más alto grado de necesidad. Por ello Dios es libre en tanto elige sin ninguna razón para elegir, porque una razón divina es una razón necesaria. La indiferencia es lo que caracteriza a la libertad de Dios, que, en tanto que creador *ex nihilo*, es capaz de transformar en libre, o sea, en voluntario, lo indiferente. No ocurre así con la voluntad humana, los agentes humanos no son libres por actuar desde la indiferencia, pero, y en esto se opone Duns Escoto a Tomás de Aquino, esto no quiere decir que sus acciones dependan de sus razones. La concepción de la libertad de Duns Escoto es más voluntarista que intelectualista. La libertad humana no reposa ni en la indiferencia de la voluntad, ni en su dependencia de las razones del agente, sino en la propia voluntad. En el pensamiento de Duns Escoto la idea del libre arbitrio de

Agustín como voluntad que es *causa sui*, logra su formulación más clara: “*Nihil aliud a voluntate est causa totalis volitionis in voluntate*” (Es la voluntad sola la causa total de la volición en la voluntad). Nuestras ideas determinan nuestras acciones, pero nosotros determinamos primero la selección de nuestras ideas. En este sentido la sola voluntad determinaría todas nuestras voliciones, en el de que cada uno de nosotros es un yo autoconstituido, un yo que es el que es por su sola voluntad.

El Renacimiento

En el Renacimiento los primeros humanistas retomarán el problema del libre albedrío en unos términos parecidos a aquellos en que se discutió la cuestión en la época helenística, a saber, preocupándose del carácter más bien azaroso o por el contrario determinado de la vida humana. Aunque esta discusión tendrá lugar sobre todo entre algunos humanistas italianos del siglo XV tales como Bruni o Alberti, y más tarde será retomada por Pomponazzi y por Maquiavelo, cobrará verdadera importancia en el seno de la controversia religiosa entre humanistas católicos y reformistas.

Lutero, basándose muy directamente en el apóstol Pablo y en las Escrituras, lleva la doctrina del libre albedrío de Agustín a sus últimas consecuencias, esto es, a su disolución. Se es libre sólo en virtud de la gracia divina y la libertad no es más que un acto de fe, un asentimiento ante el mandato divino. Luego, en tanto que la fe es fruto de la bondad divina, de la revelación, y no de la razón o de cualquier otra facultad humana, ser libre no es algo diferente de tener fe, lo cual es un don. Por lo tanto no existe realmente la libertad en tanto que capacidad de elegir en vistas a la acción. Si el creyente sólo se salva por serlo y sólo lo es en virtud de la predestinación divina, entonces no depende en absoluto del creyente el salvarse o no. La doctrina de Agustín, cuna del libre albedrío, llevada a sus últimas

consecuencias por un agustino con claros problemas con los matices, acaba siendo la tumba de ese mismo libre albedrío.

En *La libertad del cristiano* de 1520, Lutero limita la libertad al ámbito interno del ser humano, pero donde en Agustín cabía prácticamente todo, en Lutero apenas hay espacio más que para la fe cristiana: ser libre es tener fe en la doctrina cristiana, la acción no tiene nada que ver con la libertad y por ello el salvarse o no depende únicamente de la fe, no de las obras llevadas a cabo. “Es evidente que [al cristiano] nada que sea externo –llámese como se llame– puede justificarle y hacerle libre, porque su bondad y su libertad, al igual que su malicia y su cautiverio, no son realidades corporales y externas.”⁵² Los mandamientos divinos sólo tienen una función reveladora de la infame condición del ser humano, sirven para humillar al cristiano que se ve intrínsecamente incapaz de cumplirlos y le convencen de que, en efecto, nada hay que dependa de él: “Ello nos deja ver que al cristiano le basta con la fe; no necesita obra alguna para ser justificado. Si no precisa de obras, ha de tener la seguridad de que está desligado de todos los preceptos y leyes; y si está desligado, indudablemente es libre. Esta es la libertad cristiana: la fe sola.”⁵³ El cristiano ha de saber que todo cuanto intente hacer para salvarse es inútil: o ha sido agraciado o no lo ha sido, esa es toda la libertad posible. “Puesto que las obras a nadie justifican y el hombre tiene que estar justificado antes de realizarlas, resulta evidente que sólo la fe que procede de la pura gracia por Cristo y su palabra es la que justifica suficientemente a la persona y la salva (...).”⁵⁴ Sin duda no hay nada más contrario a lo que defiende Agustín, y sin embargo es inconcebible que Lutero hubiese llegado a esas conclusiones sin partir del pensamiento agustiniano. Lutero sería pues el máximo exponente de un incompatibilismo que podría llamarse *fatalista*, pues lo que hace es declarar

⁵² Lutero, *La libertad del cristiano*, 3, en 2001: 157.

⁵³ Ibidem: 160.

⁵⁴ Ibidem: 166.

que libertad y determinismo son incompatibles, pero en vez de dar cuenta de una teoría de la libertad no determinista afirma que es cierto el determinismo y que por lo tanto no somos libres.

Estas ideas hallarán su contrapunto en 1524 en el *Sobre el libre arbitrio* de Erasmo de Rotterdam⁵⁵, no en balde gran conocedor también del pensamiento de Agustín pero, como buen humanista, mucho más preocupado que el líder de la Reforma por la suerte de los seres humanos. El interés filosófico de la obra de Erasmo es mayor y, desde luego, sus argumentos más convincentes que los de la obra de Lutero, cuya principal baza argumentativa es el recurso a la autoridad de la biblia y de las epístolas de Pablo. Erasmo considera que el hombre es libre en el sentido de que no está predeterminado, y que por ello es responsable de sus actos y susceptible de labrarse por sí mismo su salvación: “¿Y por qué yo he de hacerte ahora el elogio de nuestra libertad? ¿No te parece que es una magnífica libertad vivir fuera del alcance de la fortuna y no tenerla como muestra el ceño, ni deseirla cuando abre la sonrisa, ni descorazonarse cuando se presenta adversa, ni engreírse cuando se presenta fausta? Ninguna jurisdicción tiene sobre nosotros.”⁵⁶

Lutero contestará en 1525 a Erasmo con su *Esclavo arbitrio*, en el que básicamente recupera lo ya expuesto en anteriores escritos si bien con una mayor cantidad de insultos y de invectivas contra su antagonista que, desde el punto de vista moral pero también intelectual, resulta claramente vencedor en la contienda. A modo de resumen de las ideas que desarrolla Lutero en su *Esclavo arbitrio*, sirva esta cita que recoge Schopenhauer en su *Sobre la libertad de la voluntad*:

La razón natural tiene que conceder que Dios tiene que ser de tal manera que, en virtud de su libertad, nos someta a nosotros a la necesidad. [...] Porque si se admite la

⁵⁵ Aunque Erasmo se ocupa de la libertad, más brevemente eso sí, en otras muchas obras.

⁵⁶ Erasmo, *Del menosprecio del mundo*, en 1956: 582.

omnisciencia y omnipresencia, se sigue de modo natural e incontrovertible que nosotros no nos hemos hecho a nosotros mismos o vivimos o hacemos algo por nosotros mismos, sino sólo por su omnipotencia. [...] La omnisciencia y omnipotencia de Dios se hallan en diametral contradicción con la libertad de nuestra voluntad. [...] Todos los hombres se sienten obligados a reconocer con irremisible consecuencia que nosotros no llegamos a ser lo que somos por nuestra voluntad, sino por necesidad; que, por consiguiente, no hacemos lo que nos place en virtud de una libertad de la voluntad, sino a tenor de lo que Dios ha previsto y ejecuta mediante una firmeza y consejo infalibles e inmutables.⁵⁷

En fin, el problema que todos los filósofos medievales se ocuparon de resolver lo deja como irresoluble Lutero; y tal cual lo hereda Calvino y lo radicaliza, pues a la idea de providencia luterana le suma Calvino la de predestinación. Según Calvino Dios no es sólo la causa primera omnisciente y omnipotente de todo lo que hagan los seres humanos, sino aquél que dispone y determina el destino de todos y cada uno de los hombres. Si en Lutero la salvación dependía de la fe, y ésta a su vez de la gracia divina, en Calvino la predeterminación de la salvación es más inmediata y no necesita de la gracia ni de la fe: se tiene o no se tiene, se ha sido creado para salvarse o para condenarse. Por otro lado, en el bando de los erasmistas destaca un español, Luis de Molina, que logrará hacer la gran síntesis de las tesis católicas en su obra de 1588 *La concordia del libre albedrío con el don de la gracia, la divina presciencia, providencia, predestinación y reprobación* y se convertirá en obra de referencia sobre la libertad para toda la filosofía posterior. No está exento de cierta ironía que dicho escrito, a pesar de haber sido creado con el ánimo de defender a la Iglesia frente a los ataques de Lutero (o al hombre frente a los ataques a su libertad), estuviese a punto de ser incluido en el Índice. En ella Molina defiende una visión incompatibilista de la libertad basada en la voluntad que, para ser libre, ha de ser autónoma e independiente:

⁵⁷ Schopenhauer, *Sobre la libertad de la voluntad*, 2000: 121.

Ahora bien, las potencias libres, como la voluntad angélica o humana, aunque sean potencias para querer y no querer, no están sin embargo determinadas por su propia naturaleza a tales acciones, como si estando presentes y concurrentes todos los requisitos para querer o no querer, tenga que elegir tales acciones por necesidad de su naturaleza, no pudiendo, por consiguiente, evitarlas. Es más: su libertad, el dominio de sus actos, el que sean capaces de la virtud y del vicio, del mérito y del demérito, de la alabanza y del vituperio, del premio y del castigo, consiste, precisamente, en que, aún estando presentes y concurrentes todos los requisitos para obrar, pueda evitar o no elegir cualquier acto, o bien elegir sobre uno y el mismo objeto el quererlo más que el no quererlo o, al contrario, el no quererlo más que el quererlo.⁵⁸

Lo verdaderamente novedoso de la posición de Molina, y que le valió un enfrentamiento con el jesuita Domingo Báñez y la condena papal para su libro, es que para salvar esta libertad humana de querer o no querer (independiente incluso del entendimiento, según Molina) y hacerla compatible con la gracia, la presciencia y la predestinación divinas, en lugar de redefinir la libertad, Molina redefine los poderes divinos de predecir y predeterminar. En efecto, dado que según Molina el hombre es libre de querer o no querer “aún estando presentes y concurrentes todos los requisitos para obrar”, sus acciones son imprevisibles y novedosas, totalmente incondicionadas, *independientes* de la gracia y la presciencia divinas. Pero, naturalmente, lo que no pueden es contradecir el poder y saber de Dios, por lo que Molina decide que éste, al dejar actuar de manera independiente a sus creaturas, no hace sino ser consecuente con la naturaleza que les dió, lo que, puesto que fue él quien les dió esa naturaleza, viene a significar ser consecuente consigo mismo. En cierto sentido, el destino, para no contradecir sus designios, para que no haya una acción que vaya en su contra, se ajusta a las decisiones del agente libre, que

⁵⁸ Luis de Molina, *Concordia liberii arbitrii cum gratiae donis*, VII, disp. 1, en Ocaña 1995: 64.

actúa plenamente incondicionado, aún estando predestinado. Se trata de un nuevo modelo de Dios, el “no beligerante”:

De todo lo dicho, creo que queda suficientemente claro el modo cómo la libertad de nuestro albedrío y la contingencia de las cosas se armonizan con la presciencia divina. Pues como las cosas que proceden de nuestro albedrío, o de él dependen, no son precisamente futuras por el hecho de que Dios las conozca previamente así futuras, sino que, al contrario, Dios las conoce previamente de éste o aquel modo futuras, precisamente porque así van a ser futuras por la libertad del arbitrio; y que si hubieran de ser futuras sobre lo opuesto, como muy bien podría ser, también sobre lo opuesto, -y no sobre lo que en realidad ahora se conoce-, habría recaído ese conocimiento previo desde la eternidad [...]. De donde se sigue claramente que la presciencia de Dios (...) [que] capta con absoluta certeza qué harán las causas libres, puestas en diferente orden de cosas, -siendo así que, si quisieran, bien podrían hacer aún lo contrario-, ninguna predisposición se aduciría en absoluto a la libertad del albedrío (...).⁵⁹

En fin, en el Renacimiento el debate determinismo-libertad será subsumido por la lucha (una lucha encarnizada, y cuantiosa en volúmenes de metafísica) entre reformistas y contrarreformistas, y se alejará cada vez más del terreno de la razón filosófica para acercarse al de la letra de los escritos bíblicos (una notable excepción a esta norma es la de Luis de Molina, el último gran eslabón de esa polémica). Descartes traerá consigo el nacimiento de la filosofía moderna y nuevas esperanzas para ese debate.

La filosofía moderna

El debate moderno acerca de la libertad de la voluntad está marcado por cómo lo reformula Descartes, en cuya obra las Escrituras dejan por fin de ser un argumento a favor de una determinada versión de la libertad. La visión

⁵⁹ Ibidem, ‘Apéndice’, respuesta a la objeción 2, en Ocaña 1995: 85.

que Descartes tiene de la libertad humana es muy similar a la de Duns Escoto salvo en lo que respecta a su voluntarismo. Así, Descartes también distingue la voluntad humana de la divina en tanto que la primera no se basa en la indiferencia, como sí lo hace la segunda; pero lo que no considera acertado el filósofo francés es la idea de que la voluntad humana es libre en virtud de ser voluntad de sí misma. Muy al contrario, si la libertad humana es diferente de la divina es porque su voluntad no se basta a sí misma y necesita del conocimiento, lo cual según Descartes explica el error: “(...) mis errores (los cuales por sí solos atestiguan que hay en mí imperfección), encuentro que dependen del concurso de dos causas, a saber, de la potencia de conocer que está en mí, y de la potencia de elegir, o bien de mi libre arbitrio: es decir, de mi entendimiento y conjunto de mi voluntad.”⁶⁰ En Dios el conocimiento es perfecto, lo cual le permite que su voluntad sea a su vez perfecta, con lo que Dios es tanto más libre en cuanto que elige desde la indiferencia, pues así, por su acto de voluntad que necesariamente será adecuado a la verdad, se verá determinado sólo por sí mismo. Es cierto que en el ser humano la voluntad aparece tan infinita como en Dios: “Tampoco me puedo quejar de que Dios no me haya dado un libre arbitrio, o una voluntad suficientemente amplia y perfecta, puesto que en efecto la experimento tan vaga y tan extendida, que no está encerrada entre ningunos polos.”⁶¹ Pero no puede decirse lo mismo de su conocimiento, lo que hace que el ser humano sea libre cuando su voluntad se ve determinada y no cuando es indiferente, pues una voluntad indiferente en el ser humano es sinónimo de imposibilidad de decidirse, esto es, de duda: “(...) esta indiferencia que siento, cuando no soy llevado hacia un lado más bien que hacia otro por el peso de ninguna razón, es el grado más bajo de la libertad, y hace más bien presentarse a un defecto en el conocimiento que una perfección en la voluntad; ya que si conociese siempre claramente lo que es verdadero y lo

⁶⁰ Descartes, *Meditaciones metafísicas*, “Cuarta Meditación”, 1996: 86.

⁶¹ *Ibidem*: 87.

que es bueno, no me vería nunca en el brete de deliberar qué juicio y qué elección debería hacer; y así sería enteramente libre sin ser nunca indiferente.”⁶² Si el hombre hubiese de ser libre en razón de su indiferencia, entonces el acto máximo de libertad sería el del asno de Buridán, que muere de hambre por ser incapaz de decidirse por uno de los dos montones de paja que tiene a idéntica distancia del punto en que se encuentra. Pero no, el único ser capaz de no quedarse inactivo a pesar de actuar desde la indiferencia es Dios: “En cuanto a la facultad del libre arbitrio, es obvio que la razón o la esencia de la que está en Dios es bien diferente de la que está en nosotros, tanto más en cuanto que repugna que la voluntad de Dios no haya sido durante toda la eternidad indiferente a todas las cosas que han sido hechas o que se harán jamás, no habiendo ninguna idea que represente el bien o lo verdadero, lo que hay que creer, lo que hay que hacer o lo que hay que omitir, que podamos fingir haya sido el objeto del entendimiento divino antes de que su naturaleza haya sido constituida tal por la determinación de su voluntad.”⁶³ Así, Descartes defiende que hay dos tipos de libertad: una absoluta que aparece como el clásico *liberi arbitrii indifferentiae*, y que es propia de Dios; y otra relativa que es de un tipo compatibilista, propia del ser humano, según la cual un agente es tanto más libre cuanto más determinado está por su razón a elegir correctamente y a no permanecer indiferente ante dos elecciones posibles:

Ya que, para que yo sea libre, no es necesario que sea indiferente a elegir uno u otro de dos contrarios; sino más bien, que cuanto más me inclino por uno, sea porque conozca evidentemente que el bien y lo verdadero se hallan en él, sea porque Dios disponga así el interior de mi pensamiento, tanto más libremente lo elijo y lo abrazo. Y ciertamente la gracia divina y el conocimiento natural, bien lejos de disminuir mi libertad, la aumentan más bien, y la fortifican.⁶⁴

⁶² Ibidem: 88-89.

⁶³ Ibidem, “Sextas objeciones y respuestas”, 1996: 279.

⁶⁴ Ibidem, “Cuarta Meditación”, 1996: 88.

A este análisis de la libertad en sus *Meditaciones*, Hobbes le hizo la siguiente objeción: “(...) hay que notar que la libertad como libre arbitrio es supuesta sin ser probada, a pesar de que esa suposición sea contraria a la opinión de los calvinistas.”⁶⁵ La respuesta de Descartes a Hobbes fue la siguiente: “No he supuesto o avanzado nada, en lo tocante a la libertad, más que aquello que sentimos todos los días en nosotros mismos, y que es muy conocido por la luz natural (...).”⁶⁶ Así, parece que Descartes termina sustituyendo la autoridad de la biblia por la autoridad de la autoconciencia (lo cual es muy cartesiano): la prueba última de que existe libre albedrío es que yo me siento ser libre. Varios siglos más tarde Schopenhauer se ocupará directamente del problema de si la autoconciencia nos dice o no algo acerca de nuestra libertad, pero no es momento todavía de acercarnos a él; volvamos a Hobbes.

Si la teoría de la libertad de Descartes es internista, la teoría de Hobbes es, en principio, externista. De hecho, en la primera parte del *Leviatán*, Hobbes utiliza como sinónimos “movimiento animal” y “movimiento voluntario”, opuestos al crecimiento o “movimiento vital”, con lo que la voluntad aparece claramente ligada al funcionamiento adaptativo de los seres humanos, a su cuerpo y sus necesidades naturales. Los movimientos voluntarios “dependen siempre de un pensamiento procedente de *adónde, cómo y qué*”⁶⁷, que produce un conato que puede ser un apetito o deseo, o una aversión. Los conatos son “esos pequeños comienzos del movimiento en el cuerpo del hombre, antes de que aparezcan otras acciones visibles”⁶⁸, y son la causa de la acción: si se trata de un deseo la acción tenderá *hacia* algo, si se trata de una aversión tenderá a *apartarse* de algo. Es posible que en nuestra mente tenga lugar una

⁶⁵ Ibidem, “Terceras objeciones y respuestas”, 1996: 195.

⁶⁶ Ibidem.

⁶⁷ Hobbes, *Leviatán*, 1992: 49.

concatenación de deseos y aversiones luchando entre sí antes de que al fin llevemos a cabo una acción concreta, en eso consiste según Hobbes la deliberación (que no sería exclusiva de los seres humanos, sino propia también de los demás animales), y ésta “recibe este nombre porque consiste en un poner fin a la *libertad* que teníamos de hacer o de no hacer, según nuestro apetito o aversión.”⁶⁹ Así, la voluntad es lo contrario de la libertad, o cuando menos cabe decir que la voluntad se da cuando termina la libertad, ya que la libertad es *facultad* de querer y la voluntad *acto* de querer. La libertad sólo acaece mientras oscilamos entre ciertos deseos y ciertas aversiones, pero termina cuando la deliberación transforma un deseo o una aversión posible en un hecho; entonces tiene lugar una acción voluntaria, pero no libre, pues está determinada por un deseo o una aversión concretos. “Por lo tanto, *la voluntad es el último apetito en el proceso deliberador*. Y aunque en el lenguaje común decimos que un hombre tuvo alguna vez la voluntad de hacer algo que, sin embargo, se abstuvo luego de hacer, eso no fue, propiamente hablando, otra cosa que una inclinación, la cual no produce ninguna acción voluntaria, ya que dicha acción no depende sino de la última inclinación o apetito.”⁷⁰ ¿Y en qué se basa Hobbes para hacer esta afirmación? Más arriba señalé que la teoría de Hobbes es externista *en principio*, y así es, pues Hobbes recurre, al igual que Descartes, a una experiencia interna para fundamentar su explicación de la acción libre: “(...) no puede ser ofrecida otra prueba que la experiencia propia de cada hombre, reflexionando sobre sí mismo (...).”⁷¹ Así Hobbes, que reprochaba a Descartes presuponer la existencia del libre arbitrio cuando precisamente ese era el problema, acaba haciendo lo mismo, eso sí, para demostrar precisamente lo contrario: su inexistencia. En efecto, según Hobbes toda acción es causada por una volición que es idéntica a un conato que efectivamente produce una

⁶⁸ Ibidem: 50.

⁶⁹ Ibidem: 56.

⁷⁰ Ibidem: 57.

acción, determinándola. Y según Hobbes los deseos y las aversiones no causan de cualquier manera, sino necesariamente. En su ensayo de 1684, *De la libertad y la necesidad*, defiende que todas las causas de la acción son causas suficientes, y que toda causa suficiente ha de ser también necesaria y producir necesariamente su efecto, pues si no fuese capaz de hacerlo querría decir, según Hobbes, que no era suficiente. Por ello todo cuanto es producido lo es necesariamente y esto incluye a las acciones voluntarias, pues los apetitos y aversiones son causas suficientes, de lo que resulta que “esa *definición* ordinaria de un *agente libre*, a saber, *que un agente libre es aquel que, cuando están presentes todas las cosas que son necesarias para producir el efecto, puede no obstante no producirlo*, implica una contradicción, y es un sinsentido; siendo tanto como decir que una causa puede ser *suficiente*, esto es *necesaria*, y aún así el *efecto* podría no seguirse de ella.”⁷² Hobbes representa pues el extremo que más inquieta a los pensadores libertaristas, pues no se conforma con defender una teoría compatibilista de la libertad, sino que apuesta por una versión fatalista de ese compatibilismo: no sólo nuestras acciones están determinadas, sino que lo están necesariamente. La razón de que Hobbes adopte esta postura tan radical está en su defensa del anglicanismo, lo que le obliga a leer la biblia literalmente, a la manera de Lutero, y a defender tesis similares a las de éste. Para Hobbes “(...) la *inconveniencia* de negar la *necesidad* sería que destruye tanto los *decretos* como la *presciencia* de Dios Todopoderoso (...).”⁷³ Así pues, el argumento último del autor del Leviatán, contra Descartes, sería la biblia, según la cual el ser humano no puede poseer libre albedrío, pues ello iría en contra de la omnipotencia y omnisciencia divinas. Para Hobbes “(...) [*La libertad es la ausencia de todos los impedimentos de la acción que no están contenidos en la naturaleza y cualidad*

⁷¹ Hobbes, *Of Liberty and Necessity*, en Hobbes 1966: 275.

⁷² Ibidem.

⁷³ Ibidem: 278.

intrínseca del agente”⁷⁴, esto es, para Hobbes sólo existiría la libertad negativa, de acción, no la positiva, de elección.

Podría parecer paradójico que Leibniz⁷⁵ pueda estar de acuerdo con esa definición, y que sin embargo su teoría de la libertad deba leerse como una teoría compatibilista no fatalista, como una teoría en que se defiende que existe el libre albedrío, aun cuando existe el determinismo. En sus *Ensayos de teodicea sobre la bondad de Dios, la libertad del hombre y el origen del mal* de 1710, Leibniz escribe un apéndice cuyo encabezamiento reza: “Reflexiones sobre la obra que M. Hobbes ha publicado en inglés, de la libertad, de la necesidad y del azar”. En dicho apéndice Leibniz reprocha a Hobbes que convierta a Dios en causa tanto de lo bueno como de lo malo, y que hable de necesidad absoluta donde basta la hipotética (de esta distinción me ocuparé en I.3.1.). Por lo demás, Leibniz suscribe los argumentos que Hobbes esgrime contra el incompatibilismo y contra versiones agustinistas del libre albedrío, ahora, dado que Leibniz distingue entre tipos de necesidad, puede defender una versión de la libertad conciliable con el determinismo. La interpretación de Leibniz de la determinación de las acciones del agente por sus razones es similar a la que adoptó Crisipo para hacerle frente a las críticas de fatalismo que los epicúreos dirigieron a los estoicos durante generaciones, una interpretación según la cual actuaríamos determinados pero no con necesidad, en sentido absoluto: “(...) la necesidad no debe oponerse a la volición, sino a la contingencia (...) y la necesidad no debe ser confundida con la determinación (...). Y si no nos damos cuenta siempre de la razón que nos determina o por la cual nos determinamos, es porque somos (...) [in]capaces de darnos cuenta de todo el

⁷⁴ Ibidem: 273.

⁷⁵ De acuerdo con el orden histórico aquí probablemente debería poner “Spinoza” y no “Leibniz”, no obstante, dado que Leibniz escribió específicamente sobre Hobbes, he preferido poner a aquel a continuación de éste.

juego de nuestro espíritu y de sus pensamientos (...).”⁷⁶ Así, la teoría de Leibniz es muy similar a la de Hobbes en lo que respecta a su descripción de la acción y de la determinación de la misma por parte del agente a través de la voluntad, pero no en lo que atañe a sus consecuencias. Si se obvia este aspecto, resulta sorprendente hasta qué punto las palabras que vienen a continuación recuerdan a la teoría hobbesiana de los deseos y las aversiones que causan la acción: “(...) diría que la volición es el esfuerzo o la tendencia (conatus) de ir hacia lo que encontramos bueno y lejos de lo que encontramos malo, de suerte que esta tendencia resulta inmediatamente de la apercepción que tenemos de ella; y el corolario de esta definición es este célebre axioma: que del querer y del poder, reunidos juntos, se sigue la acción, puesto que de toda tendencia se sigue la acción cuando nada se lo impide.”⁷⁷

Una vez más este punto se ve justificado por un sentimiento interno, aunque ya Leibniz comienza a hacer depender su definición de la libertad de la cohesión de su sistema: la libertad del ser humano encaja con la armonía preestablecida, pues está determinada, pero existe, pues Dios ha creado el mejor de los mundos posibles y la libertad es una perfección que le es característica. En cualquier caso, el autor que renunciará definitivamente al internismo como fundamento de la libertad y la hará depender del ordenamiento mismo del mundo, o más bien dará una explicación de la misma conciliable con dicho ordenamiento, será Baruch Spinoza, del que me voy a ocupar tras esta cita que resume de forma muy completa la teoría del libre albedrío de Leibniz:

Pero la libertad del espíritu, opuesta a la necesidad, mira la voluntad desnuda, en tanto que es distinguida del entendimiento. Es lo que llamamos el libre arbitrio y consiste en que queremos que las razones o impresiones más

⁷⁶ Leibniz, *Nuevos Ensayos sobre el Entendimiento Humano*, Libro segundo, Cap. XXI, pf. 13, en Leibniz 1966: 253.

⁷⁷ Ibidem, pf. 5: 251.

fuertes, que el entendimiento presenta a la voluntad, no impidan que el acto de voluntad sea contingente, y no le den ninguna necesidad absoluta y por así decirlo metafísica. Y es en este sentido que tengo costumbre de decir, que el entendimiento puede determinar la voluntad, siguiendo la prevalencia de las percepciones y razones de una manera, que precisamente porque, es cierta e infalible, inclina sin necesitar.⁷⁸

No está exento de alguna ironía que el autor que la mayor parte de los filósofos tienen por el máximo defensor del fatalismo haya sido precisamente el que más importancia ha concedido en sus escritos al problema de la libertad. Pocos autores, exceptuando tal vez a Agustín, Kant y Sartre, han pensado la libertad con el empeño con el que Spinoza se ha ocupado de ella. Puede decirse que, en cierta medida, Spinoza recoge el testigo de los estoicos y de Boecio y que es el máximo eslabón de una cadena que se podría prolongar con Leibniz, Hume, Schopenhauer y Wittgenstein (si bien este último autor no se ocupó extensamente del problema) hasta llegar en nuestros días a Donald Davidson y Daniel Dennett, no obstante haber muchos más de los que me ocuparé detalladamente más adelante.

Como he dicho, Spinoza suele ser tenido por el más gran enemigo de la libertad y el representante más brillante de un mecanicismo absolutamente fatalista. Tal honor se debe sin duda, entre otras cosas, a la proposición XLVIII de la parte segunda de su *Ética*: “*No hay en el alma ninguna voluntad absoluta o libre, sino que el alma es determinada a querer esto o aquello por una causa, que también es determinada por otra, y ésta a su vez por otra, y así hasta el infinito.*”⁷⁹ En esta proposición la palabra “absoluta” no carece de importancia. Spinoza se declara seguidor de Descartes, y como tal distingue entre una voluntad absoluta y una voluntad que produce acciones que están determinadas por ella

⁷⁸ Ibidem, pf. 8: 252.

⁷⁹ Spinoza, *Op. Cit.*, 1988a: 176. La cursiva es mía.

y por el entendimiento (lo cual no quiere decir que la voluntad equivalga al entendimiento), y en eso consistiría la libertad de los hombres: “ni yo ni Descartes hemos dicho jamás que pertenece a nuestra naturaleza el contener nuestra voluntad dentro de los límites del entendimiento, sino tan sólo que Dios nos ha dado un entendimiento determinado y una voluntad indeterminada (...) y que, además, esta voluntad indeterminada y perfecta no sólo nos hace más perfectos, sino que, además (...) ella nos es muy necesaria.”⁸⁰ Spinoza no entiende por “indeterminado” la ausencia de una causa necesaria, ni lo contrario de “determinado”, sino aquello que no tiene término, o sea, que es infinito. Así, la voluntad es “indeterminada” en ese sentido y “determinada” en otro, y por ello es libre: “usted dice que yo hago a los hombres tan dependientes de Dios que los convierto en algo similar a los elementos, a las hierbas y a las piedras (...). (...) [Pero] si usted hubiese captado por el entendimiento puro qué significa depender de Dios, no pensaría jamás que las cosas, en cuanto dependen de Dios, son muertas, corpóreas e imperfectas (...), sino que comprendería que, precisamente por eso, en cuanto dependen de Dios, las cosas son perfectas.”⁸¹ En efecto, para que la voluntad sea indeterminada, perfecta, ha de estar determinada por Dios, esa es la única forma de que la voluntad humana no esté dominada por algo diferente de sí misma. Lo que es necesario por depender de Dios no es necesario en el sentido de opuesto a libre, sino determinado por sí mismo, esto es, según la definición VII de la parte primera de la *Ética*, precisamente libre: “Se llama *libre* a aquella cosa que existe en virtud de la sola necesidad de su naturaleza y es determinada por sí sola a obrar; y *necesaria*, o mejor *compelida*, a la que es determinada por otra cosa a existir y operar, de cierta y determinada manera.”⁸² Si se recuerda lo que dice Spinoza en la proposición XLVIII de la

⁸⁰ Spinoza, *Correspondencia*, [129], Carta 21, 1988b: 197.

⁸¹ Ibidem [131]: 198.

⁸² Spinoza, *Op. Cit.*, 1988a: 47.

parte segunda de su *Ética*, en el alma no hay voluntad libre porque aquella se ve determinada por una cadena causal infinita. Esa cadena causal infinita acaba conduciendo a Dios, que está al principio de esa cadena causal, lo cual viene a ser decir que la voluntad del alma coincide con la voluntad divina, así, en tanto el alma participa de la sustancia divina, está actuando “en virtud de la sola necesidad de su naturaleza” lo cual, según la definición que vista más arriba, es actuar de forma libre. En este punto se separa Spinoza de Descartes, según el cual no puede considerarse que la libertad humana y la divina sean iguales. Pero en la teoría de Spinoza sólo cabe un tipo de libertad, la divina, pero esta libertad es transferida por la sustancia divina a sus atributos, como se deduce de las proposiciones XXV, XXVI, XXVII y XVIII de la primera parte de su *Ética*:

*Dios no es sólo la causa eficiente de la existencia de las cosas, sino también su esencia. [...] Una cosa que ha sido determinada a obrar algo, lo ha sido necesariamente por Dios; y la que no lo ha sido por Dios, no puede determinarse a sí misma a obrar. [...] Una cosa que ha sido determinada por Dios a obrar algo, no puede convertirse a sí misma en indeterminada. [...] Ninguna cosa singular, o sea, ninguna cosa que es finita y tiene una existencia determinada, puede existir, ni ser determinada a obrar, si no es determinada a existir y obrar por otra causa, que es también finita y tiene una existencia determinada (...) y así hasta el infinito.*⁸³

Así, el ser humano es libre porque se determina a sí mismo porque Dios, ese dios que incluye todas las cosas, se está siempre autodeterminando a sí mismo necesariamente. Por ello, la libertad humana no puede ser una libertad de indiferencia a pesar de ser indeterminada, porque por la necesidad de la esencia divina la voluntad humana no puede evitar estar determinada por las ideas que causan las acciones. A este respecto véase la proposición XLIX de la parte segunda de la *Ética*: “En el alma no se da ninguna volición, en el sentido de

afirmación y negación, aparte de aquella que está implícita en la idea en cuanto que es idea.”⁸⁴ Esta proposición tiene un corolario absolutamente revolucionario, que constituye en sí mismo una versión nueva del concepto de libertad de la voluntad: “La voluntad y el entendimiento son uno y lo mismo.”⁸⁵ No se trata ya aquí de un intelectualismo aristotélico según el cual para que una decisión sea libre ésta ha de ser fruto de una deliberación, sino de la defensa de la absoluta coincidencia del entendimiento y la voluntad, esto es, *ser libre es lo mismo que ser autoconsciente*: darse cuenta de que la necesidad acaece en uno mismo, eso es ser libre.

A modo de resumen de las ideas de Spinoza presento aquí un extracto de una de sus cartas a Ehrenfried Walther von Tschirnhaus. En ella se ve que, a pesar de lo novedoso de su enfoque, Spinoza acaba recurriendo, como Descartes, como Hobbes, como Leibniz, al testimonio de la autoconciencia para justificar la libertad:

El estado de la cuestión [del libre albedrío] abarca tres puntos. *Primero*, si tenemos algún poder absoluto sobre las cosas exteriores. La respuesta es negativa. Por ejemplo, que yo escriba ahora esta carta no está totalmente en mi poder (...). *Segundo*, si tenemos poder absoluto sobre los movimientos de nuestro cuerpo, los cuales surgen cuando la voluntad los determina. Respondo: con una limitación, a saber, que gozamos de un cuerpo sano (...). *Tercero*, si cuando me es permitido hacer uso de mi razón, puedo hacerlo libérrimamente, es decir, de forma absoluta. Y la respuesta es afirmativa. Pues ¿quién me negaría, sin contradecir su propia conciencia, que no puedo pensar, en mis pensamientos, que yo puedo escribir o no escribir? Y lo mismo se diga de la acción, puesto que lo permiten las causas externas (...).

Estoy de acuerdo con usted en que se dan causas que me determinan a ello (...). También afirmo como cierto, conforme a mi conciencia, con Descartes, que tales cosas no

⁸³ Ibidem: 79-81.

⁸⁴ Ibidem: 177.

⁸⁵ Ibidem: 179.

me fuerzan y que, por eso mismo, puedo realmente (...) omitirlo, pese a esas razones.⁸⁶

La teoría de Locke hereda el materialismo con que Hobbes elabora la suya, pero prescinde de su aparato religioso, con lo que se la puede considerar un buen ejemplo de teoría empirista de la libertad, que, naturalmente, no es concebida como libre arbitrio aunque, eso sí, también se basa en el testimonio de la autoconciencia: “Todo el mundo, creo, encuentra en sí mismo un poder para comenzar o abstenerse de, continuar o llevar a término, varias acciones en sí mismo. De la consideración de la extensión de ese poder de la mente sobre las acciones del hombre, que cada uno encuentra en sí mismo, surgen las ideas de *libertad* y *necesidad*.”⁸⁷

Según Locke, no tiene sentido preguntarse por la libertad del querer, la libertad de la voluntad, pues no le corresponde a ella ser o no libre, sino a los agentes. Así, la pregunta relevante es cuándo un agente es libre para querer, porque de hecho la voluntad nunca puede ser libre ya que lo que un agente quiere determina que elegirá hacer, y no está en poder del agente no querer lo que quiere. Así, “[l]a idea de libertad es la idea de un poder dentro de cualquier agente para hacer o abstenerse de hacer una acción concreta, de acuerdo a la determinación o pensamiento de la mente, por la cual una de ellas es preferida a otra; donde si alguna de ellas no está en poder del agente, para ser producida por él de acuerdo a su volición, entonces éste no es libre.”⁸⁸ Así, un agente es libre en tanto es capaz, lo cual no siempre es posible, de tener diferentes pensamientos motivadores de su acción⁸⁹, porque mientras es capaz de

⁸⁶ Spinoza, *Correspondencia*, [263], Carta 57, 1988b: 333-334.

⁸⁷ Locke, *Un ensayo sobre el entendimiento humano*, Libro II, Cap. XXI, 1998: 159.

⁸⁸ Ibidem: 159-160.

⁸⁹ Contra esta forma de concebir la libertad he elaborado un pequeño experimento mental, “el ejemplo del asesino poco ocurrente”, que expongo en I.2.2.

cambiar de un pensamiento a otro, es capaz de elegir lo que determinará su acción. Así resume Locke su teoría:

Primero, (...) la libertad consiste en nuestro ser capaces de actuar, o de no actuar, de acuerdo a lo queelijamos o queramos. Segundo (...) la volición, o el querer, es un acto de la mente que dirige su pensamiento a la producción de una acción, y que así ejerce su poder para producirla. [...] Tercero, dado que la voluntad no es sino una poder de la mente para dirigir las facultades operativas de un hombre para moverse o reposar, en tanto dependen de dicha dirección: a la pregunta ‘Qué es lo que determina la voluntad?’ la verdadera y adecuada respuesta es, ‘La mente’.⁹⁰

Así Locke, como Spinoza, está subordinando la libertad y la voluntad al entendimiento, pero en ningún caso llega a sugerir que la voluntad y el entendimiento sean lo mismo. Más bien la teoría de Locke es de un intelectualismo de viejo cuño. También, la teoría de Locke se parece a la de Hobbes, sin embargo hay una clara diferencia entre ambas (al margen de la ausencia de la biblia en la teoría de Locke), a saber: según Locke el deseo que mueve a la voluntad no es un conato actualizado, sino el resultado de una cierta “incomodidad”: “¿Qué es lo que determina la voluntad en lo que respecta a nuestras acciones? (...) [U]na cierta incomodidad (...) que aflige al hombre. [...] A esta incomodidad la podemos llamar *deseo*, que es una incomodidad del espíritu debido a la necesidad de un bien ausente.”⁹¹ En último término, ese bien ausente es el “bien positivo más grande”, al cual tendemos todos, que es la felicidad. Por tanto, ir retirando incomodidades, esto es, ir satisfaciendo deseos, es tender hacia la felicidad, es encauzar nuestra voluntad hacia el máximo bien. Así, al fin la felicidad es lo que determina nuestra voluntad y la libertad es nuestra facultad mental para elegir entre

⁹⁰ Locke, *Un ensayo sobre el entendimiento humano*, Libro II, Cap. XXI, 1998: 162.

⁹¹ Ibidem: 164-165.

pensamientos que nos permitan acometer acciones que deseemos. En fin, la teoría de Locke viene a ser una lectura aristotélica de la libertad.

Entrando ya de lleno en la Ilustración, en Francia ésta dejará sobre todo un interesante debate entre Bossuet y Voltaire, que si bien se centra principalmente en el sentido de la historia, alcanza también al problema del libre albedrío. Así, en tanto que Bossuet, en su *Discurso sobre la historia universal*, defiende que todo cuanto sucede, sucede según un plan de la providencia, Voltaire, por el contrario, como buen ilustrado, alberga la esperanza de que el destino de la humanidad esté en manos de sus miembros, y supone que la historia es producto de las acciones libres de los hombres. La idea de que el sentido de la historia viene determinado por la providencia parece ser algo relativamente común entre los filósofos de la historia, así Vico, como Bossuet, también defiende que la historia está regida por la providencia, pero su forma de comprender la providencia es diferente de la de Voltaire. Vico no cree que la historia sea un producto acabado a priori por los designios divinos o de la razón universal: para Vico la providencia es algo así como el reino de los fines, un proyecto ideal que actúa en la historia a través de la mano de los hombres. Así, según Vico la historia es al mismo tiempo predeterminada por Dios pero construida libremente por los hombres. Volviendo a Francia, del mismo modo que Vico tan sólo se ocupa de la libertad tangencialmente, Bossuet sólo trata dicho concepto en razón del destino de las naciones en un marco histórico. Voltaire, en cambio, si le dedicará algunas páginas en diversas obras al problema de la libertad de la voluntad propiamente dicho: así lo hace en su *Tratado de metafísica*⁹², en *El filósofo ignorante*, en *El principio de acción* y en su *Diccionario filosófico*. En esta obra le dedica un artículo a la libertad. Esas páginas son probablemente lo más ameno que se ha escrito sobre el problema que nos

⁹² A modo de anécdota quisiera señalar que en este tratado Voltaire defiende lo contrario de lo que defenderá en las demás obras.

ocupa, pero son también escasas, lo que conlleva que no tenga mucho sentido resumir aquí dicho artículo. Por ello, además de recoger a modo de muestra la cita que puede leerse un poco más abajo, tan sólo diré que Voltaire defiende que ha de entenderse la libertad como libertad negativa, como libertad de acción. Así, defiende una postura similar a la de Hobbes pero negando explícitamente que exista cualquier tipo de providencia o de necesidad divina. Y ahora, la cita prometida:

Sobre la libertad:

“B.–Pero, de una vez por todas, ¿yo soy libre?

A.–Vuestra libertad no es libre, pero vuestras acciones lo son. Sois libre de hacer cuanto tenéis el poder de hacer.

B.–Pero todos los libros que leído sobre la libertad de indiferencia...

A.–Son necedades; no existe libertad de indiferencia; es una palabra sin sentido, inventada por gentes que no lo tenían.”⁹³

De todos modos, y a pesar de estas líneas de Voltaire, los ilustrados franceses, preocupados principalmente por la libertad política y los derechos civiles, no tratarán en profundidad el problema del libre albedrío, una cuestión demasiado metafísica para los *philosophes*.⁹⁴

En Gran Bretaña, si bien también prevalecerán las cuestiones de orden práctico entre los ilustrados, la figura de Locke como precursor traerá consigo un intenso debate sobre cuestiones de fundamentación del conocimiento y de fundamentación de la moral, donde entra el problema del libre albedrío. Entre los autores que tratarán el tema en profundidad hay que destacar a dos filósofos escoceses: Thomas Reid y David Hume.

⁹³ Voltaire, *Diccionario filosófico*, 1985: 366.

⁹⁴ Aunque tal vez se pueda considerar, por ejemplo, que el problema del que se ocupa Rousseau de la oposición entre intereses privados y voluntad general sea el correlato político del problema del

En lo que respecta a Reid, en los últimos años viene despertando gran interés en muchos filósofos analíticos de corte libertarista, y las referencias a su teoría de la libertad se han hecho moneda corriente en el debate actual. A Reid se le suele considerar padre fundador de la teoría del agente causante y, dado que me ocuparé extensamente de la misma en *I.1.3.*, no me voy a detener aquí en este autor, tan sólo señalaré su rasgo fundamental: los agentes son causas de sus acciones, pero no en el mismo sentido en que lo son otras causas, la agencia conlleva un tipo especial de causación que no es equiparable a la causación mecánica que tiene lugar en la naturaleza.

En cuanto a Hume, éste sigue los pasos de Hobbes y, aunque por razones muy distintas, defiende también que la libertad no existe. En el ser compatibilista o incompatibilista entran dos cosas: una teoría acerca del funcionamiento causal del mundo y de la acción humana, y otra teoría acerca del concepto de libertad. Si resulta que la primera postula el determinismo, pero la segunda es de tipo incompatibilista, entonces uno ha de concluir que la libertad, entendida como libre arbitrio, no existe, y eso es lo que hace Hume: “(...) la libertad, cuando se oponga a la necesidad y no a coerción, será lo mismo que el azar, el cual se reconoce universalmente que no existe” pues “[s]e acepta universalmente que nada existe sin una causa de su existencia, y el azar, cuando se examina rigurosamente, no es más que una palabra negativa y no significa un poder real que esté en la naturaleza.”⁹⁵ Según Hume a las acciones humanas es necesario aplicarles la ley de la causalidad, lo cual hace que los motivos hayan de ser entendidos como causa de la acción y que un mismo motivo siempre haya de llevar consigo una misma acción. Y no sólo eso, según Hume los filósofos han creado artificiosamente una oposición entre necesidad y libertad, pero en la vida cotidiana no existe esa oposición, pues la

libre albedrío. En este sentido las aportaciones de los filósofos de la época al debate serían cuantiosas.

⁹⁵ Hume, *Investigación sobre el conocimiento humano*, Sección 8, 1980: 120 y 119.

gente suele dar explicaciones de carácter causal a las acciones propias y ajenas: “De esta manera resulta que la unión de los motivos y acciones voluntarias no sólo es tan regular y uniforme como lo es la de la causa y efecto en cualquier región de la naturaleza, sino también que la humanidad unánimemente ha reconocido esta conjunción regular, y jamás ha sido objeto de discusión ni en la filosofía ni en la vida común.”⁹⁶ Así, Hume procede con respecto a este problema wittgensteinianamente antes de Wittgenstein: el problema de la libertad es un falso problema, como el de la existencia del mundo externo, si uno lo piensa con atención se da cuenta de que no hay nadie que defienda seriamente que la libertad y la necesidad son incompatibles, sino que todo el mundo presupone lo contrario.

Pero además Hume lleva el empirismo a sus últimas consecuencias: “una serie de causas aglutinadas entre sí por lo que nos place llamar necesidad *física*. La experiencia repetida de la unión produce el mismo efecto sobre la mente, aunque los objetos unidos sean motivos, voliciones y acciones, o figura y movimiento. Podemos cambiar el nombre de las cosas, pero su naturaleza y acción sobre la mente nunca cambian.”⁹⁷ Para Kant, no será tanto la idea de que la voluntad causa lo que representará un problema, sino precisamente la pretensión de Hume de que la voluntad causa en el mismo sentido en que un antecedente causa su efecto en la naturaleza, porque Kant sí persigue la existencia de una libertad como libre arbitrio (lo cual le acerca a Reid). Pero Hume en este sentido no hace ninguna concesión: “Quizá nos cueste inicialmente fijar tan estrechos límites al entendimiento humano, pero después no encontraremos dificultad cuando apliquemos esta doctrina a las acciones de la voluntad. Pues como es evidente que éstas tienen una conjunción regular con motivos, circunstancias y caracteres, y como siempre sacamos inferencias de las unas a las otras, estamos obligados a reconocer de palabra la necesidad

⁹⁶ Ibidem: 112.

⁹⁷ Ibidem: 115.

que ya hemos admitido en todas las deliberaciones de nuestras vidas y en todos los pasos de nuestra conducta y comportamiento.”⁹⁸ Por ello Hume concluye, como Locke, que “sólo podemos entender por libertad *el poder de actuar o no actuar de acuerdo con las determinaciones de la voluntad (...)*.”⁹⁹

También, a sus argumentos a favor de su propia teoría, Hume le suma otros que pretenden ser objeciones definitivas a dos formas incompatibilistas de entender la libertad: la epicúrea y la agustiniana. Según Hume, vincular las acciones al azar para evitar caer en el determinismo trae consigo la anomia moral que a muchos filósofos les hace combatir la idea de que la libertad y la necesidad son compatibles: “según el principio que niega la necesidad y consecuentemente las causas, un hombre es tan puro e intachable tras haber cometido el más horrendo crimen, como lo era al nacer, ni su carácter se ve en manera alguna afectado por sus acciones, puesto que no se derivan de él, y la maldad de las unas nunca podrá emplearse como prueba de la depravación del otro.”¹⁰⁰ En cuanto a la idea agustiniana de la libertad:

No es posible explicar claramente cómo Dios puede ser la causa mediata de todas las acciones de los hombres, sin ser el autor de sus pecados y de su bajeza moral. Se trata de misterios que la mera mente natural, sin otra asistencia, no es capaz de tratar adecuadamente y, cualquiera que sea el sistema al que se acoja, ha de verse sumida en dificultades inextricables, e incluso en contradicciones, a cada paso que dé con respecto a tales temas. Hasta ahora se ha visto que excede todo el poder de la filosofía el reconciliar la indiferencia y contingencia de las acciones humanas con la previsión, o defender el carácter absoluto de los decretos divinos, y, sin embargo, librar a la Deidad de ser autora del pecado.¹⁰¹

Schelling volverá a ocuparse del problema del mal cuando trate la libertad, e intentará dar una respuesta satisfactoria a las dudas de Hume. A su

⁹⁸ Ibidem: 117-118.

⁹⁹ Ibidem: 119.

¹⁰⁰ Ibidem: 122.

vez la respuesta de Schopenhauer a Schelling será devastadora. Pero aún no hemos llegado al idealismo, aún resta estudiar a un autor ilustrado.

Los logros de la Ilustración alemana acerca del problema de la libertad de la voluntad se pueden resumir en la obra de Kant, uno de los filósofos que más espacio le han dedicado a ese problema, que ocupa además un lugar central en su sistema. Se puede decir que la exposición más clara del problema libertad/determinismo que en la historia de la filosofía ha habido es la tercera antinomia de la razón pura, esto es:

Tesis

La causalidad según leyes de la naturaleza no es la única de la que pueden derivar los fenómenos todos del mundo. Para explicar éstos nos hace falta otra causalidad por libertad.

Antítesis

No hay libertad. Todo cuanto sucede en el mundo se desarrolla exclusivamente según leyes de la naturaleza.¹⁰² (*KrV*, B 472-B 473)

La tesis se justifica del siguiente modo: si todo sucediese tan sólo según las leyes causales naturales tendríamos una cadena causal infinita, una sucesión de acontecimientos que, por definición serían siempre consecuencia de una causa anterior; pero ninguno de ellos podría ser una causa primera, “[a]hora bien, la ley de la naturaleza consiste precisamente en que nada sucede sin una causa suficientemente determinada *a priori*. Así pues, la proposición según la cual toda causalidad es sólo posible según leyes de la naturaleza se contradice a sí misma en su universalidad ilimitada.”¹⁰³ (*KrV*, B 474)

La antítesis se prueba así: si suponemos que hay un acto libre, entendido como fruto de “una *absoluta espontaneidad* causal que inicie *por sí misma* una serie de fenómenos”¹⁰⁴ (*KrV*, B 474), al comienzo de la cadena

¹⁰¹ Ibidem: 127.

¹⁰² Kant, *Crítica de la razón pura*, 1978: 407

¹⁰³ Ibidem: 408.

¹⁰⁴ Ibidem.

causal del conjunto de los sucesos naturales del universo comenzará un hecho causal, una determinación que crea la causalidad pero que es a su vez indeterminada, “[a]hora bien, todo comienzo de acción supone un estado anterior propio de la causa que todavía no actúa, y un primer comienzo dinámico de acción supone un estado que no está unido por ningún vínculo causal con el anterior estado de la misma causa, es decir, no se sigue en modo alguno de ese estado anterior. Así pues, la libertad trascendental se opone a la ley de causalidad.”¹⁰⁵ (*KrV*, B 474)

Así, Kant se inclina por presuponer una libertad trascendental, pero es consciente del problema que, como acabamos de ver, señala Hume: es la causalidad lo que nos permite explicar las acciones y la “libertad (independencia) respecto de las leyes de esta naturaleza nos *libera* de la *coacción* de las reglas, pero también del *bilo conductor* que todas ellas representan.”¹⁰⁶ (*KrV*, B 475) Por otra parte, Kant admite que el concepto de libertad que solemos tener, el “psicológico” (o sea, el que se le aparece a nuestra autoconciencia) es “empírico” en gran parte y por ello es difícilmente conciliable con una idea trascendental de libertad. Sin embargo, Kant defiende que en tanto esta facultad de iniciar por uno mismo una cadena causal espontáneamente es un *a priori* que estructura la acción, del mismo modo que la ley de causalidad estructura la percepción gracias las formas puras de la sensibilidad, “[n]o es tan necesario el que podamos decir cómo sea posible semejante cosa (...) aunque no entendamos en modo alguno cómo es posible que la existencia de una cosa sea puesta por la existencia de otra y aunque debamos, por ello mismo, atenernos tan sólo a la experiencia.”¹⁰⁷ (*KrV*, B 476) Así, no es la experiencia de la libertad lo que podría fundamentar la libertad

¹⁰⁵ Ibidem.

¹⁰⁶ Ibidem.

¹⁰⁷ Ibidem: 410.

trascendental¹⁰⁸, pues ésta es condición de posibilidad de la primera, “[o]curre, pues, aquí lo que se descubre siempre en el conflicto de una razón que se aventura más allá de los límites de la experiencia posible: que el problema no es *fisiológico*, sino *trascendental*.”¹⁰⁹ (*KrV*, B 563) Eso sí, un problema trascendental que también es moral, pues según Kant la libertad moral es posible porque, a diferencia del animal, que posee un *arbitrium brutum*, el ser humano posee “*arbitrium sensitivum*, pero no *brutum*, sino *liberum*, ya que la sensibilidad no determina su acción de modo necesario, sino que el hombre goza de la capacidad de determinarse espontáneamente a sí mismo con independencia de la imposición de los impulsos sensitivos.”¹¹⁰ (*KrV*, B 562) Así, la libertad moral incluye la libertad trascendental, que a su vez no puede ser fundamentada por la experiencia, lo cual nos permite entrever que la solución definitiva a la tercera antinomia ha de estar en el quiasmo entre la razón pura y la razón práctica. Aunque Kant apunta un primer esbozo de solución (que en el fondo plantea más dificultades de las que resuelve, pues se asienta en lo que Davidson ha llamado con acierto el “tercer dogma del empirismo”), pues dado que los fenómenos no son cosas en sí mismas, sino representaciones, podemos considerar que, en tanto que cosas de la naturaleza, los fenómenos se ven sometidos a la ley de causalidad, pero en tanto que representaciones son determinados libremente: “(...) podemos considerar el efecto como libre con respecto a su causa inteligible, pero, con respecto a los fenómenos, podemos tomarlo, a la vez, por resultado de esos mismos fenómenos según la necesidad de la naturaleza.”

Así, Kant está proponiendo una solución aparentemente compatibilista, pues considera que determinismo y libertad conviven en buena entente; pero realmente no es así, porque lo que ha hecho Kant es escindir el mundo y

¹⁰⁸ Esto mismo defiende Schopenhauer, digno y fiel discípulo de Kant, aunque el triunfo del idealismo en su época le hurtase ese título.

¹⁰⁹ Kant, *Crítica de la razón pura*, 1978: 465.

poner a sus dos mitades a discurrir de forma paralela, y no reconciliarlo explicando una única realidad, la del mundo, con dos prismas que son uno solo mirado desde dos puntos de vista.¹¹¹ En la *Crítica de la razón práctica* es donde puede verse con más claridad hasta qué punto Kant no logra resolver la tercera antinomia, lo cual no es extraño porque, por otra parte, en la forma en que la plantea la antinomia no puede ser resuelta:

En la antinomia de la razón pura especulativa se encuentra un antagonismo (...) entre necesidad natural y libertad en la causalidad de los acontecimientos del mundo. Quedó disuelta al probarse que no suponía un auténtico antagonismo, / cuando se considera sólo como manifestaciones fenoménicas (tal y como debe hacerse) a esos acontecimientos e incluso al propio mundo donde suceden; porque uno y el mismo agente tiene *en cuanto fenómeno* (incluso ante su propio sentido interno) una causalidad en el mundo sensible que siempre resulta conforme al mecanismo de la naturaleza, pero con respecto a ese mismo acontecimiento, mientras se considere al mismo tiempo a la persona que actúa en cuanto *noúmeno* (como inteligencia pura en su existencia no determinable conforme al tiempo), puede entrañar un fundamento para determinar esa causalidad según leyes naturales que esté él mismo libre de toda ley natural.¹¹² (*AkV*, 114)

En definitiva, la solución de Kant al problema de la libertad de la voluntad viene a ser señalar que a la *res cogitans* le compete actuar libremente y

¹¹⁰ Ibidem: 464.

¹¹¹ El corazón de muchos kantianos puede estar sufriendo al leer frases como esta: ciertamente cabe, y tal vez se debe, leer a Kant como aquel que reconcilia, gracias a la idea de mundo moral, naturaleza y voluntad (véase la *Crítica de la Razón pura*, Kant 1978: 631-632, [*KrV*, B835-836]). Sin embargo, tanto Fichte como Schopenhauer (los principales intérpretes y seguidores inmediatos de Kant) construirán sus propias teorías partiendo de lo que consideran una irrebasable escisión, propia de la teoría kantiana, entre mundo causal y moral. Ello explicaría también la vuelta de estos filósofos a Spinoza, que habría sido capaz de lidiar con esta quiebra de manera más satisfactoria que Kant (aunque para los idealistas la teoría de Spinoza fuese en exceso materialista). En cualquier caso, seguramente no se debe leer así a Kant en nuestros días, sino que conviene hacerlo *à la* Rawls; sin embargo, desde el punto de vista de la historia del problema del libre albedrío, es el Kant que aquí presento el que deja notar su influencia en el pensamiento posterior, y por ello es esta lectura la que es relevante desarrollar aquí. Los neokantianos pueden pues respirar tranquilos.

¹¹² Kant, *Op. Cit.*, 2000: 226.

a la *res extensa* determinísticamente. Así, Kant salva el hiato existente entre libertad y determinismo estableciendo otro: el del fenómeno y el noumenon.

Los avatares del problema de la libertad en la obra de Kant, así como el desarrollo de la cuestión por parte de los filósofos modernos, especialmente Spinoza y Leibniz, ocuparán el centro de las preocupaciones del idealismo alemán, cuya posibilidad misma como proyecto dependería de la resolución de la tercera antinomia, esto es, de la reconciliación de la Naturaleza y el Espíritu, o sea, de la necesidad y la libertad. El primero en transitar esta complicada senda en estos términos será Fichte, y no de manera accesoria: su sistema bien podría recibir el nombre de “idealismo de la libertad”.¹¹³ Fichte parte del hecho de la razón práctica, que según Kant a su vez instituye como postulado la libertad, pero camina en sentido inverso: el hecho primero que nos es desvelado por una intuición inmediata, y que por lo tanto es la primera piedra del edificio, que viene dado en la autoconciencia (*es* la autoconciencia), es la libertad. Pero la libertad está en abierta contradicción con la necesidad que parece ser el devenir del mundo que nos es desvelado por las representaciones. ¿Cómo reconciliar ambos extremos? En la ley moral: el imperativo categórico reúne la libertad y la necesidad. Y el imperativo categórico participa de la misma inmediatez que la autoconciencia, es el sentimiento mismo de libertad, porque la libertad se nos aparece como un “tú debes”, como un sentimiento en que se nos revela que no podemos no actuar. Así, el fundamento de la libertad es una intuición, la primera, aquella de la que depende el resto del conocimiento, cuyas raíces según Fichte se hunden en la tierra de la razón práctica y no de la razón pura. Así, en el sistema de Fichte, la solución a la

¹¹³ Efectivamente, todo el sistema de Fichte pivota sobre el problema libertad/necesidad, por ello habrá algunos filósofos que juzgarán que el resumen que aquí hago de su sistema y de lo que él entiende por libertad es insuficiente. No me cabe la menor duda de que lo es, para exponer el concepto de libertad de Fichte me sería necesario hacer un estudio completo de su *Doctrina de la Ciencia*. Creo que resulta evidente que ello no es posible.

tercera antinomia viene a ser la misma que la kantiana con una salvedad (una salvedad importante): la libertad no es un postulado de la razón práctica, posible dada su entidad nouménica; la libertad está *ab initio*, es una intuición que precede a las formas puras del entendimiento. No es necesario recurrir a las incómodas “cosas en sí” kantianas, la libertad se despliega necesariamente en la naturaleza intuitivamente como su principio rector, es el poder de determinar el mundo sensible (donde reina la necesidad) según el mundo inteligible (donde reina la acción). En este sentido, el ataque que verterá Schopenhauer sobre la relevancia del testimonio de la autoconciencia como fundamento de la libertad, es un torpedo en la línea misma de flotación del sistema de Fichte, y puede decirse que del idealismo en general, aunque Schelling se cuidará muy mucho de apelar a la autoconciencia para explicar la libertad.

En 1809 ven la luz, en el volumen *Escritos filosóficos*, las *Investigaciones filosóficas sobre la esencia de la libertad humana y de los objetos con ella relacionados* que Schelling escribe en respuesta a ciertas acusaciones de fatalismo que algunos filósofos alemanes, entre los que cabe destacar a F. Schlegel, habían hecho recaer sobre su sistema filosófico. La base de tales críticas venía a ser más o menos la siguiente: Schelling, como Spinoza, es un panteísta y por lo tanto el sistema de Schelling, como el de Spinoza, ha de ser fatalista. Schelling escribe pues sus *Investigaciones* para demostrar que el panteísmo no sólo es conciliable con la libertad, sino condición necesaria de la misma, y que si Spinoza había llegado a una conclusión diferente (que creo no es así, pero esa era la idea dominante entonces), ello se deberá a razones diferentes del panteísmo: “(...) el panteísmo al menos no hace imposible la libertad formal; por lo tanto Spinoza debe ser fatalista por otro motivo independiente del panteísmo. El fallo de su sistema no reside en modo alguno en que sitúe las cosas *en Dios*, sino en que sean *cosas*: reside en el concepto abstracto de los seres del mundo, esto es, en el

de la propia sustancia infinita, que para él es también a su vez una cosa. Por ello, sus argumentos contra la libertad son muy deterministas y de ningún modo panteístas.”¹¹⁴ Parece que para Schelling las dificultades de Spinoza no residen en su monismo, sino en un presunto materialismo.

El problema con el que se encuentra Schelling a causa de su concepción panteísta del mundo no es tanto el de la libertad sino el agustiniano problema del mal, eso sí, con notables diferencias. Para Agustín el mal debe ser entendido como privación, no así para Schelling. Según éste el mal ha de ser entendido positivamente, o cuando menos con algún tipo de principio afirmativo; de lo contrario el mal, si hubiera de ser sólo negación, pura privación, no sería nada. Esa forma positiva de entender el mal es lo que permite a Schelling pensar la libertad humana dentro del panteísmo:

Pero que precisamente este levantamiento de la voluntad propia sea el mal, se puede aclarar de la siguiente manera: la voluntad, que sale de su sobrenaturalidad para volverse a un tiempo, como voluntad universal, particular y propia de las criaturas, tiende a subvertir la relación que existe entre los principios, (...) a usar el espíritu, que sólo recibió para el centro, fuera de este último, [...] apenas se aparta la voluntad propia del centro, que es su lugar, (...) reina entonces una mera voluntad particular que ya no es capaz de reunir a las fuerzas, como la voluntad originaria, y que, por consiguiente, tiene que aspirar a formar o componer una vida propia y singular (...) a partir del ejército sublevado de concupiscencias y apetitos (...).¹¹⁵

Para que haya libertad ha de existir la posibilidad de hacer el mal¹¹⁶, y para que el mal sea posible éste ha de ser efectivamente real y no un mero vacío de bien. El suceder de la ausencia, el darse una carencia, no es una actualización genuina de una posibilidad. El mal existe, según Schelling, como poder efectivo, y en su surgimiento como poder consiste precisamente la

¹¹⁴ Schelling, *Investigaciones filosóficas sobre la esencia de la libertad...*, 1989: 349.

¹¹⁵ *Ibidem*: 183.

libertad, la posibilidad de darse el mal convierte también en posible el bien (que de lo contrario sería necesario). No es pues necesaria la noción de gracia agustiniana, en eso y en su concepción del mal se aparta Schelling de Agustín, no obstante se puede decir que a grandes rasgos está en la línea de este pensador, dado su explícito rechazo al concepto de libertad de Spinoza. El espacio, principio de la acción libre, que en Agustín llenaba una voluntad *causa sui*, en Schelling lo ocupa la mismidad tornada espíritu que produce acciones libres cuando la emoción es el germen de la acción: “(...) en la naturaleza existen determinaciones *casuales* que sólo son explicables a partir de una emoción (...) del principio irracional u oscuro de la criatura, debidas a una mismidad activada.”¹¹⁷ El explícito rechazo de Schelling al mecanicismo es lo que le permite hablar de libertad, pues de lo contrario en ningún caso podría hablar de “determinaciones casuales” o “de principios irracionales”. En cambio, postulando un espíritu del mundo Schelling puede distinguir entre actos que suceden determinados mecánicamente y que son explicables desde la razón científica; y actos que son determinados casual y no causalmente, esto es actos libres, actos fruto del ímpetu de las fuerzas emotivas del ser humano, inexplicables a priori, impredecibles, incalculables. Schelling defiende lo que vengo llamando un incompatibilismo libertario. No quiere aceptar ni una visión azarosa de la libertad a la Epicuro (“(...) si no se puede salvar a la libertad más que mediante la total contingencia de las acciones, entonces es que no es salvable en absoluto”¹¹⁸), ni una visión determinista a la Zenón (“(...) se le opone [al azar] con gran energía el determinismo (...), que afirma la necesidad empírica de todas las acciones, debido a que cada una de éstas se encuentra determinada por medio de representaciones u otras causas que

¹¹⁶ El problema de esta idea es que acaba convirtiéndose en una petición de principio a pesar de los notables esfuerzos de Schelling por no recurrir explícitamente a la noción de *causa sui*.

¹¹⁷ Schelling, *Investigaciones filosóficas sobre la esencia de la libertad...*, 1989: 207-209. La cursiva es mía.

¹¹⁸ Ibidem: 223.

residen en el pasado (...)”¹¹⁹). Schelling es partidario de “(...) aquella superior necesidad que se encuentra igual de alejada del azar y de la obligación, esto es, de un proceso de determinación externo, y que no es en realidad más que una necesidad interna que brota por sí misma de la esencia del que actúa”¹²⁰. Ese “por sí misma” sitúa inevitablemente a Schelling al lado de Agustín.

La que, desde mi punto de vista, puede ser considerada como la mejor obra escrita específicamente acerca del problema filosófico de la libertad de la voluntad, es especialmente crítica con la obra de Schelling: “(...) nada encontramos en esas investigaciones sobre la libertad que nos pudiera ilustrar fundamentalmente sobre la misma. Era de esperar, desde que se tropieza con la definición de la libertad como «facultad de lo bueno y de lo malo», definición que no está mal para un catecismo, pero nada dice en filosofía ni puede servir para mucho. [...] En realidad, sólo una pequeña parte del tratado se ocupa de la libertad, pues su parte principal es, más bien, un extenso comunicado acerca de un dios con el que nuestro autor debe tener trabada amistad íntima, ya que nos describe hasta su nacimiento (...). El comienzo está entretejido por una serie de sofismas cuya superficialidad reconocerá todo el que no se deje imponer por la osadía del tono.”¹²¹ Por supuesto estas palabras pertenecen al máximo crítico del idealismo, Arthur Schopenhauer, y están recogidas en su *Sobre la libertad de la voluntad*, uno de los dos ensayos que componen su obra de 1941 *Los dos problemas fundamentales de la ética*. Con dicho ensayo ganó Schopenhauer en 1839 el concurso organizado por la Real Sociedad Noruega de Ciencias cuyo tema era el siguiente: “¿Se puede demostrar la libertad de la voluntad humana por la autoconciencia?”. Obviamente la respuesta de Schopenhauer es negativa, lo cual no es baladí, pues, si echamos ligeramente la vista atrás, vemos hay algo que tanto Descartes como Leibniz e incluso algunos filósofos críticos con el concepto de

¹¹⁹ Ibidem: 225.

¹²⁰ Ibidem: 225.

¹²¹ Schopenhauer, *Sobre la libertad de la voluntad*, 2000: 141-142.

libre albedrío como Hobbes admiten sin discusión: todos sentimos que escogemos lo que queremos hacer voluntariamente, y ése es el fundamento de la libertad. Schopenhauer admite que todos sentimos que podemos actuar como queremos, pero señala que ese dato es irrelevante a la hora de tratar de resolver el problema de la libertad de la voluntad, pues ésta se refiere al *poder de querer lo que se quiere*, y no solamente al poder de hacer lo que se quiere, que es lo único de lo que tenemos experiencia. *Ergo* nada nos dice la autoconciencia sobre la libertad de la voluntad, ésta debe basarse en nuestra *conciencia de las demás cosas*.

Una vez respondida la pregunta de la Real Sociedad Noruega de Ciencias, a Schopenhauer no le queda más que exponer su propia teoría de la libertad, una teoría determinista: “La *idea* se convierte en motivo, como en motivo se convierte la intuición sensible, en cuanto puede actuar sobre la voluntad. Pero todos los motivos son causas y toda causalidad implica necesidad.”¹²² Así, dado que los motivos son lo que mueve nuestra voluntad a la acción y éstos son causados necesariamente por las ideas, entonces nuestras acciones están determinadas necesariamente por nuestras ideas, que a su vez están determinadas por otras ideas y así hasta el infinito porque “del mismo modo que es imposible que una bola de billar se mueva sin que reciba antes un golpe, así también es imposible que un hombre se levante de la silla sin un motivo que le empuje o lleve; con lo que su levantarse es algo tan necesario e inevitable como el rodar de la bola después del choque. [...] Porque el hombre, como cualquier otro objeto de la experiencia, es un fenómeno que se desarrolla en el tiempo y en el espacio, y como la ley de la causalidad se aplica *a priori* a todos ellos, lo que quiere decir sin excepción, también él tendrá que estar sometido a ella.”¹²³ Así, no existen las acciones espontáneas que reclaman los partidarios de entender la libertad como el fruto de una voluntad que toma decisiones desde la nada (véase

¹²² Ibidem: 82.

¹²³ Ibidem: 93.

Agustín de Hipona), o espontáneamente (véase Anselmo de Aosta), o desde la indiferencia (véase Duns Escoto): “Toda consecuencia de una razón es consecuencia necesaria, y toda necesidad es consecuencia de una razón. [Por el contrario] [u]na vez aceptado semejante *liberii arbitrii indifferentiae* hay que afirmar la consecuencia que caracteriza a ese concepto, que es que un individuo dotado de esa libertad y colocado en las mismas circunstancias exteriores, individual y completamente determinadas, es capaz igualmente de dos acciones diametralmente opuestas.”¹²⁴ Ahora, para Schopenhauer esto no quiere decir que, como pretende Lutero, no exista la libertad de la voluntad, o que nuestras acciones estén predeterminadas (como defiende Hobbes que, aún no negando la libertad, insiste en que Dios ha predeterminado nuestras acciones, que son necesariamente producidas y conocidas de antemano por él)¹²⁵:

Si ahora, como consecuencia de lo que llevamos expuesto, eliminamos por completo la libertad de las acciones humanas reconociéndolas como sometidas absolutamente a la necesidad más rigurosa, habremos llegado al punto desde el cual podemos comprender la *verdadera libertad moral*, que es de un tipo superior.

Existe otro dato de la conciencia (...). Se trata del sentimiento completamente claro y seguro, de la *responsabilidad* por aquello que hacemos, de la *imputabilidad* de nuestras acciones, lo que descansa sobre la certeza inmovible de que somos nosotros mismos los *autores de nuestros actos*.¹²⁶

Así, Schopenhauer está enfocando el problema de la libertad desde un punto de vista normativo que será dejado de lado hasta bien entrado el siglo XX, como habrá ocasión de comprobar en el próximo apartado de esta introducción. En cualquier caso, la teoría de Schopenhauer puede ser

¹²⁴ Ibidem: 52.

¹²⁵ Esa es la principal diferencia entre la definición de la libertad de Hobbes y la de Schopenhauer. Prescindiendo de Dios, las teorías de ambos autores son muy parecidas.

¹²⁶ Schopenhauer, *Sobre la libertad de la voluntad*, 2000: 151.

considerada como la formulación más clara del compatibilismo, donde se da la “coexistencia de la rigurosa necesidad de nuestras acciones con aquella libertad testimoniada por el sentimiento de responsabilidad y por medio de la cual somos los autores de nuestros actos y éstos nos son imputables moralmente.”¹²⁷ Insisto, esto no significa que no exista la libertad: “el hombre siempre hace lo que quiere y, sin embargo, lo hace necesariamente. Lo que se explica porque él es ya lo que quiere, pues de lo que él es se sigue, necesariamente, todo lo que pueda hacer.”¹²⁸

En el polo opuesto de la filosofía alemana, Hegel, al igual que Fichte, no se ocupa del problema de la libertad en su forma tradicional, sino como contorno de su sistema. Ahora, si para Fichte la libertad era principio y final de dicho sistema, para Hegel aquella sólo surge en el momento del saber absoluto, del “espíritu concebido, en su vuelta a la inmediatez estando ahí”. Por ello, en la *Fenomenología del espíritu*, se puede descubrir no una, sino muchas definiciones de libertad, que difieren en razón del momento de la conciencia o del espíritu de que se trate. Así, por ejemplo, en la conciencia de sí la libertad se corresponde con la libertad estoica, con una libertad del pensar: “Esa libertad de la conciencia de sí, emergiendo de su manifestación consciente de sí misma en curso de la historia del espíritu, se ha, como es bien sabido, llamado *estoicismo*.”¹²⁹ O también, en lo que respecta a la razón observante, cuando se fija en la “conciencia de sí en su pureza y en su trato con la realidad efectiva externa”, entonces la libertad aparece como libertad individual en tanto no se ve coaccionada por el entorno: “Si pues, en virtud de esa libertad, la realidad efectiva es capaz de esa doble significación, entonces el mundo del individuo debe ser solamente concebido a partir del individuo mismo; y la

¹²⁷ Ibidem: 155.

¹²⁸ Ibidem: 157.

¹²⁹ Hegel, *La phénoménologie de l'esprit*, I, 1999: 169.

influencia de la realidad efectiva sobre el individuo, realidad presentada como *siendo* en sí y para sí, recibe por el hecho de ese individuo un sentido absolutamente opuesto; el individuo, o bien deja estar sin molestarlo al curso de la realidad efectiva que lo influencia, o bien lo interrumpe y lo trastorna.”¹³⁰ Parece que esta sería la definición de libertad que le convendría, probablemente, al ser humano.

Sin embargo, la libertad genuina, como dije, sólo tiene lugar en el momento del saber absoluto, pues es entonces cuando se reconcilian plenamente sujeto y objeto, y por lo tanto confluyen contingencia y necesidad, libertad y determinismo, voluntad y representación. “En el saber el espíritu [...] ha conquistado el puro elemento de su ser-ahí, el concepto. El contenido es según la *libertad* de su *ser* el Sí que se aliena o la unidad *inmediata* del saber de sí mismo. El puro movimiento de esa alienación, considerado en su contenido, constituye la *necesidad* del contenido mismo. El contenido en su diversidad es, como *determinado*, en la relación, no es en sí, y su inquietud consiste en suprimirse dialécticamente a sí mismo, donde ella es la *negatividad*. La necesidad pues o la diversidad, como son el ser libre, son igualmente el Sí, y en esa *forma* del Sí, en la cual el ser-ahí es inmediatamente pensamiento, el contenido es *concepto*.”¹³¹ En el concepto coinciden libertad y necesidad, en tanto el concepto es necesario pero también principio rector, ¿de qué?, del desarrollo, del devenir del espíritu. Ahora, siendo la historia “el devenir que se actualiza *en el saber*, el *devenir mediatizándose a sí mismo*”¹³², en último término, la verdadera libertad es la libertad histórica. En la historia el espíritu se muestra como acción, como creación, siendo sin embargo despliegue: “[la] conservación [de los espíritus], bajo el aspecto de su ser-ahí libre manifestándose en la forma de la contingencia, es la historia (...)”¹³³

¹³⁰ Ibidem: 255-256.

¹³¹ Ibidem, II: 309-310.

¹³² Ibidem: 311.

¹³³ Ibidem: 312-313.

Quedaría así completado el ascenso del idealismo, interpretado como tránsito del yo al nosotros. Un periplo que va del Yo=Yo fichteano, que conlleva que la libertad se base en la autoconciencia, a la historia hegeliana, que conlleva que la libertad se asiente en el espíritu en tanto que saber absoluto. Con ello se puede dar por terminado nuestro tránsito por el idealismo, me ocuparé a continuación de un paradigma rival: el positivismo utilitarista.

La obra en la que John Stuart Mill se ocupa con más atención del problema de la voluntad es en *La filosofía de Hamilton*, pues su conocida obra *Sobre la libertad* trata acerca de la libertad civil y no de la “así llamada ‘libertad de la voluntad’, tan desafortunadamente opuesta a la mal llamada doctrina de la necesidad filosófica”.¹³⁴ Esta frase muestra claramente que Mill pertenece al bando de los compatibilistas, concretamente habría que encuadrarle en la tradición empirista con Hobbes, Locke y Hume. En el sexto libro de su *Sistema de lógica*, en que se ocupa de los fundamentos de la moral, defiende que no es lo mismo necesidad filosófica que fatalidad, pues la fatalidad implica destino, esto es, el que una mano invisible dirija nuestro devenir, que estaría ya trazado previamente. Por contra, lo único que implica la necesidad filosófica es que nuestras acciones están causadas por nuestros motivos, que a su vez están causados por nuestros deseos o aversiones, de tal forma que, si fuésemos capaces de conocer todos los móviles de una persona, susceptibles de causar sus acciones, entonces seríamos capaces de predecir su comportamiento. Y precisamente gracias a ello existe la moral, con su sistema jurídico o cuando menos con un sistema básico de premios y castigos: gracias a que el actuar humano está causado por motivos, somos capaces de describir patrones de comportamiento deseable e incentivarlos mediante premios, y de señalar patrones de comportamiento indeseable y disuadir de tenerlos mediante

¹³⁴ Mill, *Op. Cit.*, Introducción, 1974: 59.

castigos. Nuestra moral, pues, se basa en que la libertad humana se comprenda como un actuar determinado por motivos. Esto no quiere decir que el futuro esté fijado, pero en tanto la libertad está regida por la determinación, puede ser objeto de estudio de una ciencia de la moral. Lo que caracteriza a la moral no es que escape a la ley de causalidad, sino que consiste en “un tipo de investigación cuyos resultados no se expresan en el modo indicativo, sino en el imperativo”.¹³⁵ Al fin y al cabo, adecuar medios a fines es hacer un razonamiento inductivo propio de las ciencias naturales, pero aplicado a nuestro actuar:

Por tanto, los fundamentos de toda regla de arte han de buscarse en teoremas de la ciencia. Un arte, o lo que constituye un arte, consta de reglas, conjuntamente con tantas proposiciones especulativas como sean precisas para la justificación de tales reglas. El arte completo de cualquier materia incluye una selección de tal porción de la ciencia como sea necesaria para mostrar de qué condiciones dependen los efectos que el arte se propone producir. Y el Arte, en general, consiste en las verdades de la ciencia, dispuestas del modo más conveniente para la práctica, en lugar del orden más conveniente para la especulación.¹³⁶

Así, la teoría de Mill sobre la libertad debe encuadrarse en un empirismo positivista según el cual las acciones humanas son plenamente necesarias pero aún así el ser humano posee libertad de la voluntad, pues no existe esa “coacción misteriosa” que sería la fatalidad: las acciones humanas son necesarias en virtud de las propias voliciones humanas.

La solución de Nietzsche al problema de la libertad de la voluntad es tal vez la más original de cuantas llevo estudiadas, pues invierte las teorías al uso y pone fin a la escisión que planteaba su maestro Schopenhauer, basándose a su

¹³⁵ Mill, *Un sistema de lógica*, Libro VI, Capítulo XII, en 1984: 139-140.

¹³⁶ Ibidem: 145. Mill habla de arte en sentido etimológico, como ese ámbito en el que es adecuado el uso del imperativo, en el ámbito del hacer.

vez en Kant, entre representación y voluntad; un fin que se cobra como víctima a la representación y convierte todo en voluntad. Si no hay contradicción entre libertad y necesidad, no sólo actúa voluntariamente el ser humano al hacerlo causalmente, sino también la naturaleza:

En última instancia, la cuestión es si consideramos que la voluntad es realmente algo que *actúa*, en si creemos en la causalidad de la voluntad; de ser así (y, en el fondo, creer *esto* equivale a creer en la causalidad sin más), tenemos que aventurarnos a considerar la hipótesis de que no haya más causalidad que la de la voluntad. [...] En resumen, hay que atreverse a formular la hipótesis de que siempre que constatemos la existencia de «efectos», lo que sucede es que una voluntad actúa sobre otra voluntad; de que todo acontecer mecánico, desde el momento en que en él actúa una fuerza, es justamente una fuerza de la voluntad, un efecto de la voluntad.¹³⁷

Por supuesto no es lo mismo “voluntad” que “libertad”, y menos en la obra de Nietzsche, pues esa voluntad que actúa como causalidad universal no es cualquier voluntad, sino *voluntad de poder*. Habría que considerar el mundo entero como un organismo en crecimiento cuyas partes tratan de cumplir su función vegetativa en unión con las demás. Algunas partes son más bastas, sus afectos más primarios y sus acciones más pobres, otras, como el ser humano, alcanzan un grado de sofisticación superior que permite que la voluntad se haga patente. ¿Cómo? En las promesas. La bien conocida idea de Nietzsche de que el hombre es un animal al que le es lícito hacer promesas es la clave del asunto de la libertad. La voluntad de poder actúa en la naturaleza sin previsión pero no en el hombre, porque éste es capaz de elegir algunas de sus voluntades y recordarlas: “(...) hacer promesas (...) es un activo no-querer-volver-a-liberarse, un seguir y seguir queriendo lo querido una vez, una auténtica *memoria de la voluntad*, de tal modo que entre el originario «yo quiero», «yo haré» y la auténtica descarga de la voluntad, su acto, resulta lícito interponer

tranquilamente un mundo de cosas, circunstancias e incluso actos de voluntad nuevos y extraños, sin que esa larga cadena de la voluntad salte.”¹³⁷ En efecto, para poder comprometerse el ser humano debe haber aprendido a distinguir entre aquellas cosas que dependen de él y aquellas que no, porque hacer una promesa implica que ese “mundo de cosas, circunstancias...” no alterará su acto futuro. Según Nietzsche, para lograr esto el ser humano ha debido hacerse “*calculable, regular, necesario*”, porque predecir un comportamiento futuro supone ser capaz de calcular qué cabe esperar de uno como causa, qué no es casual. En definitiva, prometer algo es decir que se hará algo porque depende de uno hacerlo, o sea, es admitir que se es responsable de hacerlo, y así es como nace la libertad o voluntad moral de la voluntad de poder. “Esta es cabalmente la larga historia de la procedencia de la *responsabilidad*. Aquella tarea de criar un animal al que le sea lícito hacer promesas incluye en sí como condición y preparación (...) la tarea más concreta de *hacer* antes al hombre, hasta cierto grado, necesario, uniforme, igual entre iguales, ajustado a regla, y, en consecuencia, calculable.”¹³⁸ Así, no sólo libertad y necesidad no son incompatibles, sino que la segunda es causa necesaria de la primera.

Eso sí, el proceso es largo, hay una diferencia entre el patán previsible del principio, y el “individuo soberano” en que termina esta evolución, pues sólo a éste, al fin, le es realmente lícito hacer promesas porque ha aprendido a ser autónomo, a ser dueño de sus actos, a ser su propio señor:

Este hombre liberado, al que realmente le *es lícito* hacer promesas, este señor de la voluntad *libre*, este soberano (...) tiene (...) su *medida de valor* (...). [...] El orgulloso conocimiento del privilegio extraordinario de la *responsabilidad*, la conciencia de esta extraña libertad, de este poder sobre sí y sobre el destino, se ha grabado en él hasta su más honda profundidad y se ha convertido en instinto, en

¹³⁷ Nietzsche, *Más allá del bien y el mal*, II parte, 36: 65.

¹³⁸ Nietzsche, *La genealogía de la moral*, Tratado Segundo, 1, 1997: 76.

¹³⁹ *Ibidem*, 2: 77.

instinto dominante: - ¿cómo llamará a este instinto dominante, suponiendo que necesite una palabra para él? Pero no hay ninguna duda: este hombre soberano lo llama su *conciencia*...¹⁴⁰

Así, en Nietzsche, como en Spinoza, la libertad de la voluntad viene a consistir en la conciencia de la necesidad con que uno actúa. Por ello, según Nietzsche, la responsabilidad es producto de la necesidad y no hay contradicción alguna entre libertad y determinismo.

En *Pragmatismo* William James, en su tercera conferencia “Algunos problemas metafísicos considerados pragmáticamente”, también aplica el modo de pensar pragmático al problema del libre albedrío. James pragmatiza la cuestión ligándola a la idea de progreso moral y afirmando que “el libre albedrío carece de significado a menos que sea una doctrina de *consuelo*.”¹⁴¹ Así, según James la creencia en el libre albedrío supone tan sólo creer en lo posible, en que se puede hacer que las cosas sean de otra forma, concretamente en que las cosas pueden ir a mejor: “El libre albedrío, pragmáticamente, significa *novedades en el mundo*, el derecho a esperar que en sus más profundos elementos como en sus más superficiales fenómenos el futuro no se repita imitando idénticamente al pasado.”¹⁴² La libertad es pues fe en el futuro, confianza en la perfectibilidad humana, “[e]l libre albedrío es, pues, una teoría cosmológica general de *promesa*, como la de lo Absoluto, la de Dios, la del Espíritu o la del Plan.”¹⁴³ Para James, en definitiva, el libre albedrío es una doctrina religiosa, y *grosso* modo queda establecida como postulado por las mismas razones por las que Kant considera que ha de postularse la existencia de un dios: porque es garantía de felicidad futura.

¹⁴⁰ Ibidem: 78-79.

¹⁴¹ James 2002: 85.

¹⁴² Ibidem: 84.

La filosofía contemporánea

Aunque Russell le dedica unas palabras a la cuestión en sus *Escritos filosóficos*, en general los primeros filósofos de la tradición “analítica” (entendiendo ésta de manera muy amplia), debido probablemente a su desconfianza *vis à vis* de la metafísica, no se preocuparon demasiado por el problema determinismo/libertad,¹⁴⁴ que por el contrario interesó sobremanera a los filósofos de la tradición fenomenológica. La filosofía anglosajona no volverá a ocuparse de la libertad de la voluntad hasta la feliz reconciliación, que hizo posible Peter Strawson, del análisis y la fenomenología por medio de la metafísica descriptiva. Curiosamente, entonces el problema empezará a tratarse de manera fecunda en los sectores analíticos y perderá peso en los de tradición fenomenológica. Pero estas consideraciones pertenecen al siguiente apartado de esta introducción; aquí me compete únicamente ocuparme de la filosofía contemporánea.

Dado el hecho de que el idealismo, por un lado, y sus máximos críticos como Schopenhauer, por otro, empezaron a poner el acento en la libertad como experiencia, y por lo tanto a desligar el problema del terreno de las potencias (esencias), para situarlo en el de la existencia, no debería parecer extraño que en la filosofía contemporánea el problema de la libertad acabase interesando especialmente en los círculos existencialistas y en sus aledaños. Así Kierkegaard, en *El concepto de la angustia*, abre la vía de la que será la forma característica de entender la problemática determinismo/libertad dentro del existencialismo, esto es, como una oposición entre esencia y existencia. La esencia sería lo propio del reino animal y de la naturaleza en general, de lo

¹⁴³ Ibidem.

¹⁴⁴ Esta afirmación ha de ser matizada. Muchas de las cuestiones que ocuparon a los filósofos analíticos de la primera mitad del siglo veinte tal vez no se vincularon de manera explícita al problema de la libertad de la voluntad, pero forman parte del mismo (así, como se verá más

estable y previsible, de lo necesario. La existencia, por el contrario, sería lo propio del ser humano, de lo contingente, de lo mudable, de lo imprevisible. De la esencia puede ocuparse la ciencia estudiando mediante leyes universales sus regularidades; de la existencia ha de hacerse cargo la historia, intentando explicar las acciones humanas sin ser jamás capaz de formular unas leyes del devenir histórico del ser humano. Así, la existencia es *posibilidad*, es un poder ser; y un poder ser es un poder hacer, o sea, libertad.

Según Kierkegaard la libertad como posibilidad, esto es, como contingencia, se deja sentir de manera privilegiada en un sentimiento: el sentimiento de angustia: “(...) la angustia es la realidad de la libertad como posibilidad ofrecida a la posibilidad.”¹⁴⁵ La forma de leer esta experiencia variará en los diferentes autores existencialistas¹⁴⁶, en el caso de Kierkegaard, la angustia es interpretada como un sentimiento de lo posible, que deja patente que lo que es podría no ser lo que es, sino otra cosa, algo que podría ser espantoso. Surge entonces un temor ante la posibilidad de una realidad alternativa tormentosa, y ese temor profundo es la angustia, que nos atenaza por miedo ante ese futuro indeterminado, que es tan sólo posible, para bien o para mal. “La escuela de la angustia es la de la posibilidad, y hay que ser instruido según ésta para serlo según la propia infinitud.”¹⁴⁷ En este sentido, según Kierkegaard la angustia es una vía para acceder a lo divino, pero se trata de una vía que no voy a estudiar aquí, porque no distraería de la que nos interesa seguir y que nos conduce a continuación a la filosofía de Bergson.

Henry Bergson considera, como Hume, que el problema libertad/determinismo es un falso problema, resultado, según él, de una mala

adelante, estudiar los enunciados contrafácticos, por ejemplo, es estudiar una parte del problema de la libertad).

¹⁴⁵ Kierkegaard, *Le concept d'angoisse*, Cap. I, pf. 5, en 1973: 144.

¹⁴⁶ Por cierto que aquí voy a leer al primer Heidegger como uno de ellos.

¹⁴⁷ Kierkegaard, *Le concept d'angoisse*, Cap. V, en 1973: 252.

comprensión del estatuto del tiempo y la conciencia. La solución de Bergson al problema es realmente muy original y coincide con la de algunos filósofos libertaristas actuales que harían bien en releer el *Ensayo sobre los datos inmediatos de la conciencia*, donde encontrarían argumentos interesantes para fundamentar la diferencia que los partidarios de la teoría del agente causante establecen entre causa natural y causa agencial. Dicha diferencia la defienden en la Ilustración Thomas Reid y poco más tarde Kant (o al menos esa es la interpretación de la solución kantiana a la tercera antinomia que he manejado aquí). Ambos autores distinguen la acción de un agente de la acción de una causa natural en razón de las diferentes esencias de ambas. Bergson, en cambio, no habla de una diferencia esencial entre la conciencia y la naturaleza, esto es, entre el agente y las cosas; sino de una diferencia en la forma de darse la temporalidad en dichos ámbitos. Así, mientras que el tiempo desde el punto de vista de las cosas es espacial, en la conciencia se manifiesta como duración. En efecto, cuando se observa una sucesión de acontecimientos naturales, no se ve realmente una sucesión de momentos hilados, pues todos ellos son externos los unos a los otros: sucede una cosa, luego otra, luego otra... y nosotros las contamos. Por ejemplo, “cuando sigo con los ojos, en la esfera de un reloj, el movimiento de la aguja que corresponde a las oscilaciones del péndulo, no mido la duración, como parece creerse; me limito a contar simultaneidades, cosa que es muy diferente. Fuera de mí, en el espacio, no hay nunca más que una única posición de la aguja y del péndulo, pues de las posiciones pasadas no queda nada.”¹⁴⁸ Por el contrario, en la conciencia el tiempo se revela como duración, los estados se suceden sin distinguirse, se *vive* el tiempo. En fin:

[E]n el espacio no hay ni duración ni siquiera sucesión, en el sentido en que la conciencia toma estas palabras: cada uno de los estados llamados sucesivos del

¹⁴⁸ Bergson, *Ensayo sobre los datos inmediatos de la conciencia*, 1999: 81.

mundo exterior existe solo, y su multiplicidad no tiene realidad más que para una conciencia capaz de conservarlos primero y de yuxtaponerlos después exteriorizando a los unos con relación a los otros. Si ella los conserva es porque esos diversos estados del mundo exterior dan lugar a hechos de conciencia que se penetran, se organizan insensiblemente juntos y ligan el pasado al presente como efecto de esa misma solidaridad.¹⁴⁹

¿En qué afecta esto a la libertad? En que ésta existe precisamente en virtud de la vivencia especial que del tiempo tiene la conciencia. El determinismo sólo puede explicar una sucesión de acontecimientos naturales, pues la causalidad necesita que haya dos sucesos, uno previo y otro posterior, *distintos* (en un sentido espacial) pero idénticos entre sí, para que uno pueda ser causa determinante del otro. Pero en la conciencia nunca se dan dos acontecimientos idénticos, porque los acontecimientos previos pasan a formar parte del acontecimiento posterior. La conciencia no puede descomponerse en múltiples “yo” que se suceden los unos a los otros, es una misma entidad en constante evolución, y por ello es impredecible, a diferencia de lo que ocurre con las cosas. Así: ni hay causas determinantes para la acción de la conciencia externas a la propia conciencia, con su historia de acontecimientos pasados y presentes, y por ello no es correcto el determinismo como forma de dar cuenta de la libertad humana; ni hay una voluntad que causa espontáneamente las acciones de la conciencia tal y como pretenden los libertaristas, y por ello tampoco es correcta la idea de libre arbitrio para explicar la libertad humana. Deterministas y libertaristas conciben la actividad de la conciencia como una sucesión de momentos externos los unos a los otros, esto es, espacialmente; pero la conciencia *dura*, se va construyendo como un todo coherente paulatinamente según va discurriendo en la vida, en el tiempo, y por ello “[s]e llama libertad a la relación del yo concreto con el acto que realiza. Esa relación

¹⁴⁹ Ibidem: 89.

es indefinible, precisamente porque somos libres.”¹⁵⁰ Así, los actos libres son producto y al mismo tiempo constructores de la conciencia, cada “yo” se define en sus actos. Para Bergson es un error pensar, como hacen Hobbes, Locke o Stuart Mill, que ciertos sentimientos causan nuestras acciones: es imposible que un deseo o una aversión sean la causa de una acción porque ninguno de ellos puede actuar por separado, la conciencia se compone de todos esos sentimientos y actúa como un todo, por eso *yo* soy libre. Sólo si nuestros actos efectivamente se corresponden con nuestra conciencia, y no con *un* estado de conciencia o con *un* aspecto de nuestra conciencia, somos libres. “En suma, somos libres cuando nuestros actos emanan de nuestra personalidad entera, cuando la expresan, cuando tienen con ella ese indefinible parecido que a veces se encuentra entre la obra y el artista.”¹⁵¹ Pero podemos no ser libres, a menudo actuamos por costumbre, por tradición, porque tenemos ciertas manías... En estos casos actuamos como cosas, somos predecibles y no se nos puede considerar libres. Así, no todas nuestras acciones son libres, lo cual permite explicar que a veces sea posible predecir el comportamiento humano. Algunos de nuestros actos están determinados, pero precisamente por ello no son libres, con lo que la libertad es algo diferente del determinismo, pero eso no la hace compatible a su vez con una voluntad libre autodeterminada. Según Bergson ni los compatibilistas ni los incompatibilistas tienen razón, el problema libertad/determinismo, tal y como decía Hume, tal y como defenderá Wittgenstein, es algo que debería ser superado:

El problema de la libertad ha nacido, pues, de un malentendido: ha sido para los modernos lo que fueron para los antiguos los sofismas de la escuela de Elea y, como esos mismos sofismas, tiene su origen en la ilusión por la que se

¹⁵⁰ Ibidem: 152.

¹⁵¹ Ibidem: 123.

confunde sucesión y simultaneidad, duración y extensión, cualidad y cantidad.¹⁵²

Se trata de una confusión en el lenguaje que sería inevitable siempre que éste tienda a dar cuenta de hechos que le son perfectamente comprensibles a la conciencia de manera inmediata, pero que son difícilmente expresables; y la “libertad es, pues, un hecho y, entre los hechos que se observan, no lo hay más claro.”¹⁵³

Heidegger reúne la experiencia de la angustia y la experiencia del tiempo, uniendo así las lecturas de la libertad de Kierkegaard y de Bergson.¹⁵⁴ En *Ser y tiempo*, Heidegger defiende que “el «ante qué» de la angustia es el mundo en cuanto tal. (...) [Q]uiere decir que los entes intramundanos carecen tan absolutamente en sí mismos de importancia, que únicamente gracias a esta insignificatividad de lo intramundano se impone el mundo en su mundanidad.”¹⁵⁵ El mundo en cuanto tal es lo posible, la mundanidad del mundo es su contingencia, ¿pero por qué ésta habría de angustiar al *dasein*? Según Heidegger porque “[a]rroja al «ser ahí» contra aquello mismo por que se angustia, su «poder ser en el mundo» propio. La angustia «singulariza» al «ser ahí» en su más peculiar «ser en el mundo», que en cuanto compresor se proyecta esencialmente sobre posibilidades. Con el «por qué» del angustiarse abre, por ende, la angustia al «ser ahí» como «ser posible» y como aquello que únicamente en sí y por sí puede ser como singularizado en la singularización.”¹⁵⁶ Así, como ya señalara Kierkegaard, en la angustia se nos revela nuestra condición de entes posibles, esto es, entes que están por hacerse

¹⁵² Ibidem: 166.

¹⁵³ Ibidem: 154.

¹⁵⁴ Aunque Heidegger llega a decir de la “exégesis del tiempo de Bergson” que es “perfectamente imprecisa e insuficiente bajo el punto de vista ontológico” (Heidegger, *El ser y el tiempo*, pf. 66, 1971: 361).

¹⁵⁵ Heidegger, *Op. Cit.*, pf. 40, 1971: 207.

¹⁵⁶ Ibidem: 208.

y de los cuales depende plenamente hacerse de una forma o de otra. En fin, la angustia nos revela que somos libres, esto es, que somos plenamente responsables de nuestro ser, pues somos un poder ser, una posibilidad que depende sólo de sí misma:

La angustia hace patente en el «ser ahí» el «*ser relativamente al* más peculiar ‘poder ser’», es decir, el *ser libre para* la libertad del elegirse y del empuñarse a sí mismo. La angustia pone al «ser ahí» ante su «*ser libre para*» (*propensio in*) la propiedad de su ser como posibilidad que él es siempre ya. Pero este ser es al par aquello a cuya responsabilidad es entregado el «ser ahí» en cuanto «ser en el mundo».¹⁵⁷

Así, la libertad del *dasein* no puede ser absoluta, pues parte de su ser-en-el-mundo, pero desde el punto de vista de la mundanidad es absolutamente incondicionada, es más, consiste en la incondicionalidad misma del *dasein*, que es un ente que está siempre por ser, una “propensión a”, por lo que, al fin, el verdadero ser del *dasein*, esto es, el ser propiamente ser “ahí” del *dasein* es la libertad.¹⁵⁸

Sartre es uno de los autores que más concienzudamente se ha ocupado del tema de la libertad, y lo ha hecho con éxito. Su teoría de la libertad es probablemente la mejor teoría libertarista¹⁵⁹ de la tradición filosófica, o cuando menos cuenta con el aliciente de ser casi la, si no *la*, única teoría libertarista

¹⁵⁷ Ibidem.

¹⁵⁸ Tal vez Heidegger pondría el grito en el cielo si pudiese ver la lectura existencialista que hago de su texto. No obstante, para el tipo de historia que estoy elaborando esta interpretación de *Ser y tiempo* es la más relevante, pues es la lectura que fue hecha mayoritariamente tras su publicación y en los decenios siguientes. Por lo tanto, *Ser y tiempo* influyó en el desarrollo del problema de la libertad entendida como obra existencialista, y aquí me interesa dicha obra sólo en tanto participó del debate libertad/determinismo.

¹⁵⁹ He de precisar aquí que, en la cuarta parte de *El ser y la nada*, Sartre elabora una teoría de la acción que completa su doctrina de la libertad, y que desde el punto de vista de la unión de ambas materias no está tan claro que Sartre sea exactamente un libertarista, pues su teoría de la acción tiene tintes compatibilistas. Prefiero considerarle un libertarista porque así viene haciéndose desde que publicara sus obras, y por otra parte ello no es completamente inexacto. Tal vez puede considerarse que Sartre es incompatibilista desde el punto de vista de la metafísica, y compatibilista desde el punto de vista de la teoría moral.

abiertamente atea. Una de las sentencias de Sartre que ha hecho fortuna es su conocido “estamos condenados a ser libres”. Esta sentencia deja claro que el problema para Sartre no es el de si somos libres o no (se da por hecho que sí), sino el de cómo lo somos y hasta qué punto es fácil o no asumir que somos seres libres y comportarnos como tales. La seguridad acerca de la condición libre del ser humano viene dada, como en el caso de Kierkegaard y el de Heidegger, por la experiencia de la angustia, una experiencia que en Sartre aparece como náusea ante la contingencia de la existencia. En efecto, cuando uno trata de pensar el porqué de que haya algo más bien que nada, atendiendo a las cosas particulares, lo que se descubre es la absoluta gratuidad de las mismas, su total y completa contingencia. Todo parece estar de más, las cosas están, sin más, no se las puede deducir necesariamente, sólo existen. El vértigo que causa este descubrimiento, en que uno se ve mareado por el exceso de existencia, por el exceso de gratuidad, es la náusea, esto es, una forma más de angustia existencial.

Pero la náusea no es únicamente algo negativo, pues permite llevar a cabo un autodescubrimiento: no sólo todo lo que me rodea es contingente, sino que yo mismo lo soy. Soy pues, posibilidad, y tengo poder para serme, esto es, soy libre: “(...) el hombre toma conciencia de su libertad en la angustia, o, si se prefiere, la angustia es el modo de ser de la libertad como conciencia de ser, y en la angustia la libertad está en su ser cuestionándose a sí misma.”¹⁶⁰ En este sentido estoy condenado a ser libre, porque nada me viene dado, me tengo que inventar a mí mismo, tengo que decidir quién quiero ser desde la nada, y se trata de algo que no puedo posponer porque hacerme ser alguien es lo mismo que vivir, que existir. Existir es irme haciendo quien soy, por ello, por así decirlo, la esencia de mi existencia es la libertad: “La libertad humana precede a la esencia del hombre y la hace posible; la esencia del ser humano

¹⁶⁰ Sartre, *El ser y la nada*, 1943: 64.

está en suspenso en su libertad. Lo que llamamos libertad es, pues, indistinguible del *ser* de la «realidad humana». El hombre no es *primeramente* para ser libre *después*: no hay diferencia entre el ser del hombre y su «*ser-libre*».”¹⁶¹

Desde el punto de vista del debate entre compatibilistas e incompatibilistas, ¿dónde queda Sartre? Como ya he dicho la suya es una teoría libertarista, Sartre no considera que mis motivos determinen mi conducta, en algún momento yo elijo ser quien soy, y la prueba de que esto es así es precisamente el sentimiento de angustia:

[S]i pudiera captarme a mí mismo en mis relaciones con esos posibles como una causa que produce sus efectos. En este caso, el efecto definido como mi posible estaría rigurosamente determinado. Pero cesaría entonces de ser *posible*: se convertiría simplemente en por-venir. Así, pues, si quisiera evitar la angustia y el vértigo, bastaría que pudiera considerar los motivos (instinto de conservación, miedo anterior, etc.) que me hacen denegar la situación encarada como *determinante* de mi conducta anterior (...): sería preciso que captase en mí un riguroso determinismo psicológico. Pero precisamente me angustio porque mis conductas no son sino *posibles*, y esto significa justamente que dichos motivos, aunque constituyendo un conjunto de motivos *para* apartar esa situación, son captados por mí al mismo tiempo como insuficientemente eficaces.¹⁶²

El caso es que para señalar algo como causa, debería ser capaz de analizar mis acciones *a posteriori*, pero la libertad se vive, por definición, *a priori*, y por ello angustia. Cuando experimento la libertad, según Sartre, yo no me siento causa de mis futuras acciones, porque a éstas las concibo como posibles, pero dado que la libertad es precisamente la medida de mi conciencia, no puede consistir en la causalidad: “(...) para evitar el miedo, que me presenta un porvenir trascendente rigurosamente determinado, me refugio en la

¹⁶¹ Ibidem: 60.

¹⁶² Ibidem: 67.

reflexión, pero ésta no tiene otra cosa que ofrecerme sino un porvenir indeterminado. Esto significa que, al constituir cierta conducta como *posible*, me doy cuenta, precisamente porque ella es *mi* posible, de que *nada* puede obligarme a mantener esa conducta.”¹⁶³ Nosotros nos apercebimos de la libertad como un acaecer indeterminado, como el suceder de lo posible. Por supuesto, yo puedo darme cuenta de que mi yo futuro es resultado de mi yo presente; pero en el presente, antes de ser quien seré, no puedo evitar no verme como no siendo quien seré, todavía no lo soy, luego tan sólo soy un posible: “(...) ciertamente, tiendo con todas mis fuerzas hacia aquel que seré dentro de un momento, al doblar ese recodo; y, en este sentido, hay ya una relación entre mi ser futuro y mi ser presente. Pero, en el seno de esta relación, se ha deslizado una nada: yo no *soy* aquel que seré.”¹⁶⁴

Por ello el ser humano, en tanto que conciencia, posee libertad de la voluntad porque es proyección, es un no ser, o cuando menos un ser que es un no ser aún: “Esa libertad que se nos descubre en la angustia puede caracterizarse por la existencia de ese nada que se insinúa entre los motivos y el acto. Mi acto no escapa a la determinación de los motivos *porque* yo sea libre, sino que, al contrario, la estructura de los motivos como ineficientes es condición de mi libertad.”¹⁶⁵ Para que mis motivos fueran la causa completa de mis acciones, deberían constituir mi yo, pero el ser humano, puesto que no tiene esencia, por definición no cuenta nunca con un yo completo, sino con un yo perpetuamente cambiante, perpetuamente autonconstituyente y nunca constituido. La existencia como esencia implica un no ser algo nunca, sino un continuo, pertinaz e inacabable serse: “la libertad que se manifiesta por la angustia se caracteriza por una obligación perpetuamente renovada de rehacer

¹⁶³ Ibidem.

¹⁶⁴ Ibidem.

¹⁶⁵ Ibidem: 70. Le he enmendado ligerísimamente la plana al traductor. El subrayado es mío.

el Yo que designa al ser libre.”¹⁶⁶ Así, el agente no determina sus acciones, son sus acciones las que, en cierto sentido, determinan al agente.

Para eliminar cualquier atisbo de incompreensión respecto a la teoría de la libertad de Sartre terminaré citando su obra divulgativa *El existencialismo es un humanismo*, donde resume brevemente dicha teoría:

Si, en efecto, la existencia precede a la esencia, no podremos explicar nunca haciendo referencia a una naturaleza humana dada y fijada; dicho de otro modo, no hay determinismo, el hombre es libre, el hombre es libertad. Si, de otra parte, Dios no existe, no encontramos frente a nosotros valores u órdenes que legitimarían nuestra conducta. Así, no tenemos ni detrás de nosotros, ni dentro de nosotros, en el dominio numinoso de los valores, justificaciones o excusas. Es lo que expresaría diciendo que el hombre está condenado a ser libre. Condenado, porque no se ha creado a sí mismo, y de otra parte no obstante libre, porque una vez arrojado al mundo, es responsable de todo lo que hace. El existencialista no cree en la potencia de la pasión. [...] Piensa que el hombre es responsable de su pasión. [...] Piensa entonces que el hombre, sin ningún apoyo y sin ningún socorro, está condenado en cada instante a inventar al hombre.¹⁶⁷

Con esta cita puede darse por terminado el espacio dedicado al existencialismo, y por poco no puede darse también por cerrado el dedicado a la historia de la filosofía en este trabajo. Tan sólo resta estudiar brevemente a Wittgenstein.

No puede considerarse que Wittgenstein se ocupase atentamente del problema de la libertad de la voluntad, sin embargo impartió unas *Clases sobre el libre albedrío* que sería injusto no mencionar aquí. Puede decirse que la postura de Wittgenstein es similar a la de Hume, pero a las ideas de este autor le suma los argumentos pragmatistas propios de su segunda época, y esto resulta ser

¹⁶⁶ Ibidem.

¹⁶⁷ Sartre, *Op. Cit.*, 1996: 39-40.

una combinación muy productiva (a la que recurriré de cuando en cuando a lo largo de este trabajo). En los apuntes de las citadas clases, Wittgenstein comienza mostrando su perplejidad ante el hecho de que se suelen oponer libertad y necesidad, y termina señalando que sin duda esto se debe al hecho de que la necesidad implica la posibilidad de que algo sea expresado mediante una ley. Pero según Wittgenstein lo que teme la gente no es que sus actos puedan ser explicados mediante una ley, sino que estén constreñidos por ella, esto es, que la leyes preexistan a nuestros actos, de tal forma que obliguen a que dichos actos sean los que son. Pero ello no es así:

La idea de que las leyes están ya escritas, y que nosotros simplemente las adivinamos. Lo que nos impulsa a proponer esta metáfora es que de hecho tenemos leyes naturales. Hay un ámbito muy grande donde no hemos encontrado leyes naturales. [...] No hay razón por la que, incluso si hubiera regularidad en las decisiones humanas, yo no habría de ser libre. No hay nada en la regularidad que haga a ninguna cosa libre o no libre. La noción de constreñimiento está presente si se piensa en la regularidad como constreñida; como producida por los raíles. Si, además de la noción de regularidad, se introduce la noción de: ‘Ha de moverse de este modo porque los raíles están puestos de este modo’.¹⁶⁸

Pero el caso es que esta visión fatalista no viene dada por la realidad, es un modo particular de mirar la realidad, según Wittgenstein un “extraño modo”. ¿En qué sentido extraño? En el de “alejado del sentido común”, entendiendo éste a la manera de Moore y sobre todo de las *Investigaciones filosóficas* y del *Sobre la certeza*. Según Wittgenstein nos inquieta la determinación cuando la vemos desde formas rebuscadas que no constituyen sino experimentos mentales pero no temores reales, como por ejemplo cuando comparamos a un ser humano libre con una persona metida en una celda y que cree moverse libremente por la habitación cuando en realidad sus

¹⁶⁸ Wittgenstein, *Op. Cit.*, 1997: 410-411.

movimientos son controlados por otra persona desde una habitación justo debajo de la suya: “Tiene un mecanismo, controlado mediante una manivela, y ustedes observan (mediante un espejo) que me muevo exactamente de la manera que él quiere que lo haga. Entonces alguien se acerca y dice ‘¿Estabas siendo arrastrado de un sitio a otro? ¿Eras libre?’, y yo digo: ‘Claro que era libre.’”¹⁶⁹ Estaríamos tentados de decirle que se equivoca, y estaríamos legitimados para decírselo, ¿pero cuándo vamos a tener ocasión de vivir una escena así, que haría relevantes nuestras dudas? “Estamos comparando el caso de un ser humano con esos casos especiales en los que *diríamos* que un hombre estaba determinado: en que diríamos que él pensaba que decidía libremente, pero que en realidad estaba constreñido. ¿Por qué habríamos de sentirnos inclinados a comparar los casos ordinarios con ese caso tan especial. // A veces, cuando busco una llave frenéticamente, pienso: ‘Si un ser omnisciente me observa en este momento, debe de estar riéndose de mí. ¡Qué gracia debo de hacerle a la deidad, viéndome buscarla sabiendo él dónde está!’ Supóngase que yo preguntara: ¿Hay alguna buena razón para mirarlo desde esta perspectiva?”¹⁷⁰

Así, mientras la mayor parte de sus contemporáneos que se ocupaban del problema de la libertad decidían recurrir a intensas experiencias internas para buscar la raíz de la libertad de la voluntad, Wittgenstein opta por una estrategia externista. Resulta que el problema de la libertad, de hecho, no es un problema real, no se trata de algo que tenga sentido en nuestro lenguaje, la libertad no es inteligible como problema, sólo como hecho:

Supóngase que dijera algo parecido a ‘Sólo soy una máquina’. Ustedes podrían preguntar: ¿qué significan estas palabras? ¿Qué quiere decir ‘cómo una máquina’? Yo diría: ‘Bien, está comprándose a sí mismo con una piedra que cae.’ En ‘¿Qué significan?’ yo incluiría ‘¿Para qué sirven sus palabras?’. [...] Supóngase que alguien que no hubiera estado nunca bajo presión dijera un buen día ‘Ahora soy libre para

¹⁶⁹ Ibidem: 414.

¹⁷⁰ Ibidem: 415.

hacer lo que quiera', ¿se está informando a sí mismo de que es libre?¹⁷¹

Y bien, con Wittgenstein doy por terminado el repaso histórico al problema del libre albedrío. En el siguiente apartado me voy a ocupar de la forma que viene tomando dicho problema en la filosofía de los últimos cuarenta años.

2. La cuestión hoy

Algunos filósofos dicen que no saben lo que es la tesis del determinismo. Otros dicen, o sugieren, que saben lo que es. Entre estos, algunos –los pesimistas tal vez- sostienen que si la tesis es cierta, entonces los conceptos de obligación moral y responsabilidad no pueden ser aplicados realmente (...). Otros –los optimistas tal vez- sostienen que esos conceptos y prácticas de ningún modo pierden su *raison d'être* si la tesis del determinismo es cierta. [...] Hay otra opinión que es expresada menos frecuentemente: la opinión, podría decirse, del genuino escéptico moral. [...] Si se me pregunta a cuál de estos bandos pertenezco, debo decir que es al primero de todos, el bando de aquellos que no saben lo que es la tesis del determinismo. Pero eso no me impide sentir alguna simpatía hacia los otros, y el deseo de reconciliarlos.¹⁷²

Como acabamos de ver, las últimas grandes discusiones en torno al concepto de libre albedrío se dieron en el ámbito del existencialismo y la fenomenología. Resulta curioso que en la última parte del siglo XX, y los primeros años del XXI, dentro de la filosofía continental apenas Paul Ricoeur se haya ocupado vagamente del problema de la libertad de la voluntad, y en cambio en el ámbito de la filosofía analítica (o filosofía postanalítica anglosajona como la llaman algunos autores) se haya convertido en una de las cuestiones más extensamente tratadas, especialmente en los treinta últimos años. Por ello aquí me centraré en las fuentes de esta tradición, aunque siempre que me sea posible acudiré a los clásicos cuyas teorías he intentado resumir más arriba, pues aún hoy y con el ingente número de publicaciones

¹⁷¹ Ibidem: 419.

¹⁷² Strawson, P. F., "Freedom and Resentment", en Fischer & Ravizza 1993: 45.

que hay al respecto, en ocasiones resulta difícil dar con argumentos más convincentes que los que puedan haber desarrollado Spinoza o Schopenhauer, por ejemplo.

Tres figuras destacan en el panorama actual, no especialmente por sus teorías, que también, sino sobre todo por su maestría al haber sido capaces de reformular los problemas que implica la discusión acerca de la libertad de la voluntad en unos términos que han hecho avanzar el debate filosófico de forma interesante, clara y fructífera. Estos tres nombres son Peter F. Strawson, Harry Frankfurt y Peter Van Inwagen. Naturalmente hay muchos otros que han elaborado teorías tal vez más interesantes incluso que las de estos tres filósofos (mis modelos serán sobre todo, además de Peter Strawson, Daniel Dennett, Robert J. Wallace, Gary Watson y Susan Wolf), no obstante ninguna de esas teorías habrá sido capaz de tratar la cuestión de la libertad de la voluntad sin tener en cuenta las tesis de Strawson, Frankfurt o Van Inwagen. A lo largo de este trabajo me referiré pues a menudo a sus teorías; no obstante aquí quiero dar un breve repaso “histórico” que permita entender el lugar que ocupan en la discusión, lo cual de paso nos permitirá saber en qué estado se encuentra hoy en día el problema de la libertad de la voluntad.

De los tres es Van Inwagen quien más ha tratado la cuestión metafísica, y su figura creo resulta indispensable porque gracias a él contamos hoy por hoy con la formulación más clara del argumento básico de los incompatibilistas en contra de los compatibilistas. Esto, aunque se buscó con ahínco a lo largo de casi toda la historia de la filosofía, tal vez por una mala definición de los términos como sugirió Hume, o tal vez porque no se contaba con las herramientas conceptuales que lo permitieran, no fue realmente posible hasta la publicación en 1981 de *An Essay on Free Will* de Van Inwagen. No me corresponde todavía exponer aquí de forma pormenorizada su argumento, pero sí señalar que toda discusión sobre el problema determinismo/libertad que se precie no puede dejar de pelearse con el

argumento de Van Inwagen, que podría resumirse del siguiente modo: todo lo que depende de acontecimientos pasados y de las leyes naturales no depende de nosotros; si el determinismo es cierto todo lo que hacemos depende de acontecimientos pasados y de las leyes naturales, luego si el determinismo es cierto lo que hacemos no depende de nosotros.

Así mismo, todo filósofo que se interese por la cuestión debe tratar el problema de las posibilidades alternativas, esto es, debe buscar una explicación al hecho de que los seres humanos, cuando llevamos a cabo una acción, tenemos la impresión de que dicha acción no es nuestra única posibilidad, sino que podríamos elegir hacer otra cosa, que para cada momento hay varios cursos de acción posibles y, aparentemente, de forma simultánea. Hoy por hoy es casi imposible tratar este problema sin discutir las ideas de Harry Frankfurt, y sin leer detenidamente su “Alternate Possibilities and Moral Responsibility” de 1969. Una vez más no se trata de que su solución sea la mejor, sino de que Frankfurt fue capaz de poner en claro las cosas realmente importantes que estaban en juego en lo que respecta al requisito de las posibilidades alternativas, base de la libertad de la voluntad. Y algo así hizo Strawson con el problema de la responsabilidad moral.

La forma que toma en la filosofía de la segunda mitad del siglo XX el problema del libre albedrío es la de una pregunta sobre el concepto de responsabilidad moral. Esta pregunta estuvo especialmente viva en los años cincuenta y sesenta del pasado siglo y durante esos años Peter F. Strawson publicó un artículo capital del cual me ocuparé extensamente en este trabajo. El artículo de Strawson supuso un cambio para la discusión acerca de la responsabilidad similar al que supuso *Individuals* para la filosofía analítica en general. De todo lo que se produjo entonces acerca del tema de la responsabilidad moral prácticamente nada ha sido capaz de conservar la

vigencia, con sus más de cuarenta años, del “Freedom and Resentment”¹⁷³ de Strawson de 1962. La influencia de ese escrito fue enorme y propició importantes cambios en la forma de tratar el problema de la libertad y el determinismo. Como dije más arriba, hoy en día es muy raro que la literatura filosófica que trata acerca del concepto de responsabilidad moral no tome en consideración las ideas de Strawson.

Antes de que Strawson publicase su artículo la discusión en torno al concepto de responsabilidad moral se centraba en la elucidación del significado de los conceptos de determinismo y libertad. Así, se subordinaba el estudio del concepto de responsabilidad moral a las diferentes concepciones del determinismo existentes. Dentro de este debate podían distinguirse claramente tres puntos de vista enfrentados:

- 1- El *incompatibilista* [o *compatibilista*] – *fatalista o escéptico*, según el cual determinismo y libertad son incompatibles [o compatibles] y, puesto que el determinismo es cierto [es falso], o bien la libertad no existe y por lo tanto no está justificado hablar de responsabilidad moral o sencillamente no tiene sentido el concepto de libertad.¹⁷⁴ Serían los seguidores de Lutero, por un lado, y de los escépticos radicales del helenismo y de Hume por el otro.
- 2- El *incompatibilista - libertarista*, según el cual determinismo y libertad son incompatibles pero el determinismo es falso, cuando menos en lo que respecta al ámbito de la acción humana, y por lo tanto existe la libertad (una libertad no-

¹⁷³ Strawson, P. F., “Op.Cit.”, *Proceedings of the British Academy* 48 (1962): 1-25. Reimpreso en Fischer & Ravizza 1993: 45-66.

causal) y está justificado hablar de responsabilidad moral.¹⁷⁵

Serían los seguidores de los epicúreos y de Agustín.

- 3- El *compatibilista*, según el cual determinismo y libertad son compatibles o incluso cabe decir que la libertad es inconcebible sin determinación, de lo cual se sigue que determinismo y libertad coexisten y podemos hablar de responsabilidad moral.¹⁷⁶ Serían los seguidores de los estoicos y de Spinoza.

A mediados del siglo veinte, cuando se formaron estos tres bandos, la postura compatibilista estaba dominada por un enfoque *consecuencialista* o *utilitarista* consistente en justificar el uso del concepto de responsabilidad moral en razón de la utilidad social de las prácticas que el uso de dicho concepto conlleva.¹⁷⁷ Básicamente el argumento utilitarista podría enunciarse de la siguiente manera: nuestras prácticas sociales de premio y castigo se basan en las ideas de libertad y responsabilidad, esas prácticas son útiles desde el punto de vista social, debemos pues concebir al ser humano como libre y responsable para seguir sacando partido de nuestras prácticas de premio y castigo. Aristóteles y Tomás de Aquino, entre otros, recurren en varias ocasiones a argumentos de este tipo. Sin ir más lejos el primero nos dice: “(...) como dijo Sócrates, ser virtuoso o vicioso es cosa que no está en nosotros [...]

¹⁷⁴ Puede consultarse como referente de esta posición el artículo de Hoppers “What Means This Freedom”, en Berofsky 1966: 26-45.

¹⁷⁵ En este caso un buen referente serían los artículos de Campbell “In Defence of Free Will”, reimpreso en Campbell 1967; y “Is ‘Freewill’ a Pseudo-problem?” en *Mind* LX 240 (1951): 446-465, reimpreso en parte en Berofsky 1966: 112-135.

¹⁷⁶ Esta postura está particularmente bien presentada en Hobart “Free Will as Involving Determination and Inconceivable Without It”, en Berofsky 1966: 63-95, y en Russell 1993, puntos 24-30.

¹⁷⁷ Buen ejemplo de la postura utilitarista clásica es el artículo de Schlick “When is a Man Responsible”, en Berofsky, 1966: 55-63. El artículo que cita Strawson en “Freedom and

Aserto tal no es cierto; porque ¿a qué se debe que el legislador prohíba las malas acciones, y anime a que se lleve a cabo las dignas y virtuosas? ¿Por qué señala penas para el que obra mal y para el que deja de obrar bien? Absurdo sería legislar sobre tales cosas si es que no dependen de nosotros.”¹⁷⁸ Y Tomás, a su vez, comete esta falacia de petición de principio: “El hombre posee libre albedrío, pues en caso contrario holgarían los consejos, las exhortaciones, los preceptos, las prohibiciones, los premios y los castigos.”¹⁷⁹ Una de las aspiraciones de Strawson en “Freedom and Resentment” fue hacer una defensa del compatibilismo que no fuera utilitarista. Puede decirse que en esto el artículo de Strawson logró un importante éxito pues los defensores actuales del compatibilismo rara vez recurren a argumentos utilitaristas. En cualquier caso han sido muchas y muy importantes las publicaciones que en los últimos años se han venido haciendo en torno a los conceptos de libertad de la voluntad y de responsabilidad. Por ello este trabajo, si bien no menospreciará las fuentes de la tradición, se hará eco principalmente de las más recientes aportaciones al debate.

Y ahora sí, tras este repaso que nos ha permitido dar con una serie de términos sin los cuales sería imposible avanzar en la discusión, puede dar comienzo el análisis del concepto de libertad de la voluntad partiendo del punto más conflictivo del debate y en que ha de basarse toda teoría acerca de la responsabilidad moral, a saber: la pregunta por la compatibilidad o incompatibilidad de determinismo y libertad.

Resentment” como referente del utilitarismo es el de Nowell-Smith “Freewill and Moral Responsibility” en *Mind*, 1948.

¹⁷⁸ Aristóteles, *Gran Ética*, Libro I, (1187a 7 y 15), en Aristóteles 1932: 24.

¹⁷⁹ Tomás de Aquino, *Summa teológica*, I p., L XXXIII, 1, en 1942: 246.

I- LA LIBERTAD DE LA VOLUNTAD

En esta primera parte trataré de mostrar que existe la libertad de la voluntad, pero no cualquier versión de la misma, sino una libertad compatible con el determinismo. Para lograrlo haré tres cosas que coinciden con los tres capítulos de esta primera parte del trabajo: (I-1.) falsar las principales estrategias incompatibilistas; (I-2.) dar una definición del requisito de las posibilidades alternativas que las haga compatibles con el determinismo; y (I-3.) distinguir entre determinación y predeterminación, así como entre indeterminación e impredecibilidad, a fin de deshacer el prejuicio que suele llevar a muchos filósofos a oponer libertad y determinismo.

Así, en el primer capítulo, tras exponer los principales problemas a los que se enfrentan las teorías de la libertad, tanto compatibilistas (*I.1.1.*) como incompatibilistas (*I.1.2.*), haré un análisis detallado de las principales teorías incompatibilistas actuales (*I.1.3.*) y terminaré encajándolas en alguno de los patrones libertaristas clásicos (*I.1.4.*), a saber: el de Epicuro, el de Agustín o el de Kant. Con esta operación trataré de mostrar que las teorías incompatibilistas de la libertad, para resolver el problema de la oposición existente entre agencia y azar, necesitan postular alguna variante de la “causa sui” que es incompatible con cualquier versión del monismo materialista que es punto de partida de este trabajo.

A ese capítulo crítico le sigue uno constructivo, en él hago un análisis fenomenológico de las posibilidades alternativas, esto es, de la sensación que generalmente todos tenemos tras hacer algo de que “podríamos haber obrado de otro modo”. Trataré de interpretar la intuición de que contamos con posibilidades alternativas de una forma que pueda dar cuenta de esa sensación tal y como se da en nuestra vida cotidiana, siendo al mismo tiempo compatible con el determinismo.

Por último me enfrentaré al principal argumento que existe en contra del compatibilismo, a saber: la idea de que si todo está determinado por las leyes de la naturaleza y por acontecimientos anteriores, entonces nada depende de nosotros. En el primer apartado del capítulo (I.3.1.) trataré de mostrar que este argumento es incorrecto porque en él se usan varios conceptos de forma incorrecta. Así, no se distingue entre posibilidad lógica y física, ni entre la indeterminación de un suceso y el hecho de que dicho suceso sea indeterminable por el observador, esto es, entre lo que es y lo que es cognoscible. Tras estas precisiones, en el último apartado de este mismo capítulo (I.3.2.) esbozaré una teoría de la personalidad, como compuesta *grasso modo* de carácter y virtud, cuyo boceto no podré terminar hasta el último capítulo del trabajo (III.4.).

Terminaré con una recapitulación de lo visto en esta primera parte.

I-1. El jardín de senderos que se bifurcan: el requisito de las posibilidades alternativas

En todas las ficciones, cada vez que un hombre se enfrenta con diversas alternativas, opta por una y elimina las otras; en la del casi inextricable Ts'ui Pên, opta –simultáneamente– por todas. *Crea*, así, diversos porvenires, diversos tiempos, que también proliferan y se bifurcan.¹⁸⁰

En 1941 Jorge Luis Borges publicó *El jardín de senderos que se bifurcan*, compilación de ocho cuentos que debe su título al último de ellos. En ese cuento Borges relata la huida de un hombre al que acaban llevando sus pasos hasta la mansión de Stephen Albert, sinólogo que lleva años intentando resolver el misterio de la obra de Ts'ui Pên, eminente político oriental que tiempo atrás decidió retirarse de la vida pública para escribir una novela y

¹⁸⁰ Borges, “El jardín de senderos que se bifurcan”, en 1971: 111-112.

construir un laberinto. Albert ha descubierto que novela y laberinto no son dos obras distintas. El laberinto no es un laberinto físico sino temporal y la propia novela es ese laberinto, un “libro que es un acervo indeciso de borradores contradictorios [...] en el tercer capítulo muere el héroe, en el cuarto está vivo.”¹⁸¹ Esas contradicciones resultan ser producto del empeño de Ts’ui Pên por elaborar una novela que diese cuenta de la realidad de su protagonista de tal forma que abarcase no sólo las elecciones efectivamente tomadas por aquel, sino también todos los demás posibles cursos de acción que podría haber elegido seguir: “De ahí las contradicciones de la novela. Fang [el protagonista de la novela], digamos, tiene un secreto; un desconocido llama a su puerta; Fang resuelve matarlo. Naturalmente, hay varios desenlaces posibles: Fang puede matar al intruso, el intruso puede matar a Fang, ambos pueden salvarse, ambos pueden morir, etcétera. En la obra de Ts’ui Pên, todos los desenlaces ocurren; cada uno es el punto de partida de otras bifurcaciones.”¹⁸²

Así, la novela de Tsui Pên está hablando de la libertad de la voluntad. Es un lugar común en filosofía considerar que ser libre consiste precisamente en tener capacidad para elegir entre diversos cursos de acción posibles... si tan sólo hubiese una posibilidad, si viviésemos en un jardín sin bifurcaciones, entonces no seríamos libres.

A esta exigencia de contar con diversos futuros entre los que elegir se le denomina en la filosofía de nuestros días el *requisito de las posibilidades alternativas*, y puede decirse que es una condición *sine qua non* de la libertad de la voluntad, por ello la teoría de la libertad que trataré de esbozar en este trabajo intentará dar cuenta de esta exigencia cuyo cumplimiento es una parte fundamental de cualquier teoría de la libre voluntad.

¹⁸¹ Ibidem: 109.

¹⁸² Ibidem: 112.

Ahora, antes de entrar de lleno en el problema de las posibilidades alternativas es necesario explicar lo que su aceptación como requisito de la libertad está presuponiendo. Intuitivamente, como ya se ha dicho, se puede comprender que contar con posibilidades alternativas supone que cada vez que vamos a actuar podemos elegir entre más de un curso de acción, pues de darse el caso contrario, esto es, que sólo nos fuese posible cada vez llevar a cabo la acción que efectivamente llevamos a cabo, entonces podríamos pensar que estamos actuando obligatoriamente (y por lo tanto sin libertad). Ahora, esta descripción de lo que puede considerarse contar con posibilidades alternativas es demasiado vaga e imprecisa. Para saber de veras de qué se está hablando es necesario, como mínimo, realizar la siguiente distinción entre dos supuestos ínsitos en la idea de las posibilidades alternativas, a saber:

- “(...) que debe haber varios senderos genuinamente disponibles para el agente (al menos algunas veces correctamente acompañados con el comportamiento del que sea cuestión).”¹⁸³
- “(...) que el agente (y no alguna fuerza exterior o la mera suerte) elige qué sendero será el sendero futuro.”¹⁸⁴

Al primero de estos supuestos lo llamaré a partir de aquí indistintamente con el nombre del todo, esto es: “requisito de *posibilidades alternativas*”.

Al segundo de estos supuestos lo llamaré a partir de aquí “el requisito del *control del agente* sobre sus acciones”.

Tradicionalmente se ha considerado que estos dos requisitos no son siempre fáciles de cumplir a la hora de elaborar una teoría acerca de la libertad,

¹⁸³ Fischer 1999: 99.

y se tiende a pensar que las teorías compatibilistas incumplen el primero de estos requisitos (pues desde el determinismo todo curso de acción aparece como necesario y por lo tanto no hay otras posibilidades diferentes de la actualizada), mientras que las incompatibilistas incumplen el segundo (pues el indeterminismo conlleva una presencia de la casualidad en la acción que elimina la determinación de la misma por parte del agente). “El problema de la libertad de la voluntad en sus trazos más gruesos es este. La libertad de la voluntad parece ser incompatible tanto con el determinismo como con el indeterminismo. La libertad de la voluntad parece, por tanto, ser imposible. Pero la libertad de la voluntad también parece existir. Lo imposible por tanto parece existir. Una solución al problema de la libertad de la voluntad sería una forma de resolver esta aparente contradicción.”¹⁸⁵

En este apartado me dedicaré a analizar los problemas a los que se enfrentan las teorías de la libertad más comunes cuando tratan de satisfacer los requisitos de las posibilidades alternativas y del control del agente. Así, en *I.1.1.* examinaré el argumento de Van Inwagen en contra de las teorías compatibilistas de la libertad, y en *I.1.2.* veré el problema al que se enfrentan el propio Van Inwagen y en general los teóricos incompatibilistas a la hora de elaborar una teoría de la libertad inmune a dicho argumento. En *I.1.3.*, estudiaré cuáles son las posibilidades que hay de resolver estos problemas, centrándome en las más recientes teorías incompatibilistas, y sus dificultades a la hora de cumplir con el requisito del control del agente. Terminaré, en *I.1.4.*, tratando de demostrar que ninguna de versiones del incompatibilismo es plenamente satisfactoria, a menos que postule entidades metafísicas extraordinarias.

¹⁸⁴ Ibidem.

¹⁸⁵ Van Inwagen 2000: 12.

De momento vamos pues a estudiar en qué sentido puede decirse que el determinismo es incompatible con la libertad de la voluntad.

I.1.1. El problema del determinismo

Es normal en filosofía que el problema de la libertad de la voluntad reciba el nombre de “problema determinismo/libertad”. Este hecho radica en la opinión de que el tipo de necesidad causal que entraña el determinismo no es compatible con las posibilidades alternativas. En la época helenística esta opinión se defendió mediante el “argumento maestro” formulado por Diodoro, referente a la verdad o falsedad de enunciados predictivos. A este argumento se enfrentaron directamente Crisipo y otros estoicos, pero a pesar de las críticas estoicas diferentes versiones del “argumento maestro” fueron apareciendo una y otra vez a lo largo de la historia de la filosofía: el “argumento de la incompatibilidad”, el “argumento de la inevitabilidad”, hasta llegar a su versión actual como “argumento de la consecuencia” (*consequence argument*). Aquí voy a recoger la versión modal de Peter Van Inwagen que es la que más a menudo se cita hoy en día.

La formulación más común del argumento de la consecuencia es como sigue:

Si el determinismo es cierto, entonces nuestros actos son la consecuencia de las leyes de la naturaleza y de sucesos del pasado remoto. Pero no depende de nosotros lo que sucedió antes de que nacióésemos, y tampoco depende de nosotros cuáles sean las leyes de la naturaleza. Por tanto la consecuencia de esos hechos (incluyendo nuestros actos) no depende de nosotros.¹⁸⁶

Se trata, como decía más arriba, de una explicitación del problema determinismo/libertad, y se lo puede rastrear en la obra de la mayor parte de los filósofos incompatibilistas que en el mundo han sido. El mérito de Van

¹⁸⁶ Van Inwagen 1983: 56.

Inwagen ha sido dar con una formulación más precisa de este argumento. Como es de suponer dicha formulación consiste en una formalización del mismo. Antes de abordarla es pues preciso establecer el significado no-formal de una serie de expresiones formales de las que hace uso Van Inwagen en su versión del argumento de la consecuencia.¹⁸⁷

Así:

- “p” es cualquier oración.
- “Np” es “p (es un hecho) y nadie tiene, ni tuvo nunca, oportunidad alguna para hacer que p (sea un hecho)”.
- “ $\Box p$ ” es “es necesario que p”, esto es, “p es verdadero en cualquier circunstancia posible”.
- “ \supset ” es “implica” o “si...entonces”.
- “&” es “y”.
- “L” es “conjunto de afirmaciones verdaderas que describen las leyes de la naturaleza”.
- “Po” es “conjunto de afirmaciones verdaderas que describen el estado del mundo en un tiempo anterior a la existencia de cualquier ser vivo”.
- “P” es “cualquier afirmación verdadera”.

Además, el argumento formal de Van Inwagen necesita de dos reglas diseñadas por él:

- Regla (α): si $\Box p$, entonces uno puede inferir Np
- Regla (β): si Np y $N(p \supset q)$, entonces uno puede inferir Nq.¹⁸⁸

¹⁸⁷ Para el desarrollo del argumento de Van Inwagen me baso en “When Is the Will Free?”, en O’Connor 1995, y en “The Incompatibility of Free Will and Determinism”, en Kane 2001, si bien su formulación original es la que aparece en Van Inwagen 1983.

¹⁸⁸ Existe otra versión de la regla (β) en Ginet, C. 1980. “The Conditional Analysis of Freedom”, en *Time and Cause; Essays Presented to Richard Taylor*, Philosophical Studies Series in philosophy, ed. Peter Van Inwagen, vol. 19, pp. 171-86. Dordrecht: D. Reidel Publishing. Lleva por nombre “modus ponens para la necesidad potencial” (*modus ponens for power necessity*).

Dado todo esto, la versión modal del argumento de la consecuencia de Van Inwagen es la siguiente:

(1)	$\Box (Po \ \& \ L) \supset P$	
(2)	$\Box (Po \supset (L \supset P))$	Lógica estándar
(3)	$N(Po \supset (L \supset P))$	Regla (α)
(4)	NPo	Premisa
(5)	$N(L \supset P)$	Regla (β)
(6)	NL	Premisa
(7)	NP	Regla (β)

Según Van Inwagen (1) constituye una definición muy precisa del determinismo, pues diríamos de algo (P) que está causalmente determinado si, como reza (1), P es implicado necesariamente (\Box) por un conjunto de afirmaciones verdaderas que describen las leyes de la naturaleza (L) y por un conjunto de afirmaciones verdaderas que describen el estado del mundo en un pasado remoto (Po).¹⁸⁹

El argumento en su conjunto puede ser traducido y adaptado de la siguiente forma (que no es la más precisa pero sí la más clara):

- (1) Nuestras leyes naturales y nuestro pasado implican que P necesariamente.
- (2) Si es nuestro pasado el que es, entonces, si son nuestras leyes naturales las que son, entonces P necesariamente.
- (3) Si es nuestro pasado el que es, entonces, si son nuestras leyes naturales las que son, entonces P sin que ello dependa de nosotros.
- (4) El pasado no depende de nosotros.

¹⁸⁹ Así, Van Inwagen está presuponiendo que es lo mismo la necesidad lógica que la determinación causal, esa es la clave de su argumento. En 1.3.1. trataré de mostrar que el precisamente por ello el argumento no funciona, de momento voy a presumir que es correcto.

- (5) No depende de nosotros el que las leyes naturales impliquen que P.
- (6) Las leyes naturales no dependen de nosotros.
- (7) P no depende de nosotros.

La conclusión (7) se sigue de las premisas (1), (4) y (6), de la aplicación de leyes elementales de la lógica y de las reglas (α) y (β). Decir esto viene a ser lo mismo que afirmar que el que nada dependa de nosotros se sigue de que el determinismo es cierto (1), el pasado no depende de nosotros (4) y las leyes naturales no dependen de nosotros (6), siempre que las reglas (α) y (β) sean correctas, cosa que por el momento no discutiré.

En principio no parece que sea posible objetar nada a (4) y a (6) luego, si queremos defender que nuestras acciones dependen de nosotros, esto es, que los seres humanos somos libres, entonces tenemos que negar la premisa (1), el determinismo.

La premisa (1) parece incompatible con las posibilidades alternativas, pues nos conduce a afirmar en (7) que NP. La existencia de posibilidades alternativas, tal y como la formula Van Inwagen, significa que hay algún hecho p tal que alguien tiene o tuvo la oportunidad de hacer que p (sea un hecho): justamente lo que niega NP. Según Van Inwagen, si aceptamos el determinismo nada de lo que es depende de nadie para ser ni para ser lo que es: depende tan sólo de las leyes naturales y de acontecimientos pasados que no caen bajo nuestro control, pero el requisito de las posibilidades alternativas consiste precisamente en el poder de hacer que algo sea o que no sea dependiendo de lo que queramos, esto es, dependiendo de nosotros.

Consideremos un ejemplo concreto para verlo más claramente. Normalmente encontramos que dependió de Truman el que cayesen sendas bombas atómicas sobre Hiroshima y Nagasaki. Si seguimos el argumento de Van Inwagen, si el determinismo es cierto, entonces, aunque Truman fue

quien tomó la decisión de utilizar la bomba atómica, *no dependió* de él: su decisión de arrojar la bomba atómica fue consecuencia de la acción de las leyes naturales a través del tiempo, de tal forma que una serie de antecedentes remotos, anteriores incluso a la existencia de Truman, causaron necesariamente su decisión. Por lo tanto no dependía de Truman el que la bomba atómica fuese o no arrojada, el no arrojlarla no era una posibilidad real.

En este sentido vemos claramente que el determinismo (tal y como lo define Van Inwagen en (1)) tiene un serio problema con las posibilidades alternativas (tal y como fueron definidas en la p. 92). Si alguien pretende defender que determinismo y posibilidades alternativas son compatibles, entonces deberá dar una definición diferente de “determinismo” o de aquello en que consiste “contar con posibilidades alternativas”.¹⁹⁰ Van Inwagen, así como los demás incompatibilistas, lo que hacen es optar por prescindir del determinismo.

Ahora, el argumento de Van Inwagen depende en gran medida de que las reglas (α) y (β) sean correctas, y ¿qué puede decirse de esas reglas? La regla (α) es una definición de “necesario” según la cual si decimos de algo que es necesario (por ejemplo que Truman ordenase bombardear Hiroshima y Nagasaki) entonces estamos diciendo que ese algo se da sin que nadie tenga posibilidad de hacer que no sea (Truman no tuvo ninguna posibilidad de hacer que la bomba atómica no cayese sobre Hiroshima y Nagasaki). El argumento de Van Inwagen para defender esta regla es una reducción al absurdo: negarla supondría que “alguien tiene elección con respecto a lo que es necesariamente verdadero” y nadie “ha sugerido nunca que los seres humanos tengan elección acerca de lo que es necesariamente verdadero”¹⁹¹. En principio no hay nada

¹⁹⁰ Se puede anticipar aquí que yo voy a elegir la segunda opción. El siguiente capítulo (I.2.) está plenamente dedicado al problema de conciliar posibilidades alternativas y determinación, lo que queda de este lo voy a dedicar a explicar por qué elegir la negación del determinismo no es una buena solución.

¹⁹¹ Van Inwagen 1975, en Kane 2001: 75.

que objetar a esta regla, sin embargo más adelante veremos cómo puede distinguirse entre diferentes tipos de necesidad y que la definición dada por la regla (α) no es válida para todos esos tipos de necesidad.

En lo que respecta a la regla (β), el propio Van Inwagen reconoce que dicha regla se fundamenta tan sólo a través de la intuición y de un argumento que no es concluyente, a saber: que no se le ocurre ningún ejemplo de (β) que pueda tener premisas verdaderas y una conclusión falsa.¹⁹² No es mi intención extenderme aquí sobre esta cuestión (quien sí tenga dicha intención hará bien en consultar los artículos mencionados en la nota 192), lo que aquí interesa es señalar que, aunque es necesario formular el argumento de la consecuencia de otro modo a fin de hacerlo plenamente correcto desde el punto de vista lógico, parece esencialmente válido e implica que determinismo y libertad son incompatibles.

“La conclusión del presente capítulo es, por tanto, que si el determinismo es cierto, entonces nadie tiene elección acerca de nada, justamente en el sentido de *tener elección* que es relevante en los debates acerca de la responsabilidad moral.”¹⁹³

¹⁹² En efecto este argumento no es concluyente, y de hecho en su artículo “A reconsideration of an Argument against Compatibilism” (McKay, T. & Johnson, D. en *Philosophical Topics* 24, 1996, pp. 113-122), Thomas McKay y David Johnson suministran un contraejemplo, no exactamente de la regla (β) pero sí de una que se deriva de ella, la regla de la “aglomeración”. No voy a entrar aquí a analizar dicho contraejemplo porque en un artículo de respuesta (Van Inwagen 2000) Van Inwagen reconoce su error pero reconstruye la regla (β) de tal forma que la hace inmune a esa crítica y en general salva su argumento redefiniendo “Np” del siguiente modo “(es un hecho que) p y toda región a la que cualquiera tenga, o haya tenido, acceso solapa a p” (obviamente esta definición se refiere a una región en el espacio lógico). Esta redefinición hace al argumento de Van Inwagen lógicamente correcto y más exacto, no obstante prefiero centrarme aquí (y así suelen proceder la mayor parte de los estudiosos del tema) en la primera versión del argumento por ser más intuitiva y estar expuesta en los términos en que viene planteándose la cuestión determinismo/libertad desde siempre (hablando de necesidad, posibilidades y de hechos que dependen o no de nosotros).

¹⁹³ Van Inwagen, en Kane 2001: 80.

I.1.2. El problema del indeterminismo

Como se ha visto el compatibilismo tiene un serio problema con la primera parte del requisito de las posibilidades alternativas, pues “si mis elecciones y acciones son consecuencia del pasado y de las leyes naturales, entonces no tengo realmente posibilidades alternativas abiertas (del tipo requerido para la responsabilidad moral)”¹⁹⁴, por ello parecería conveniente optar por una solución incompatibilista para dar cuenta de la libertad de la voluntad.

Como se pudo ver en la introducción hay al menos dos formas de ser incompatibilista: según una de ellas no existe la libertad; según otra de ellas, la del libertarista, sí existe la libertad pero es incompatible con el determinismo. El primer tipo de incompatibilista puede conformarse con esgrimir el argumento de la consecuencia y deprimirse en silencio (aunque según sus propias ideas en realidad él no estaría haciendo nada), sin embargo para el libertarista “postular simplemente el indeterminismo no disipa inmediatamente las profundas inquietudes sentadas por el argumento de la consecuencia: una respuesta de este tipo le deja a uno con el posterior reto de proporcionar una teoría libertarista positiva de la libre agencia.”¹⁹⁵

En efecto, las teorías incompatibilistas de la libertad también tienen un serio problema con la exigencia de posibilidades alternativas, sólo que en lo que respecta a la segunda parte del requisito, esto es, con el control del agente sobre sus acciones. No aceptar el determinismo supone aceptar alguna forma de configuración indeterminística del mundo que en primera instancia es posible definir de forma general del siguiente modo: “nuestras elecciones y nuestro comportamiento son genuinamente aleatorios.”¹⁹⁶

¹⁹⁴ Fischer 1999: 100.

¹⁹⁵ Fischer & Ravizza 1993: 11.

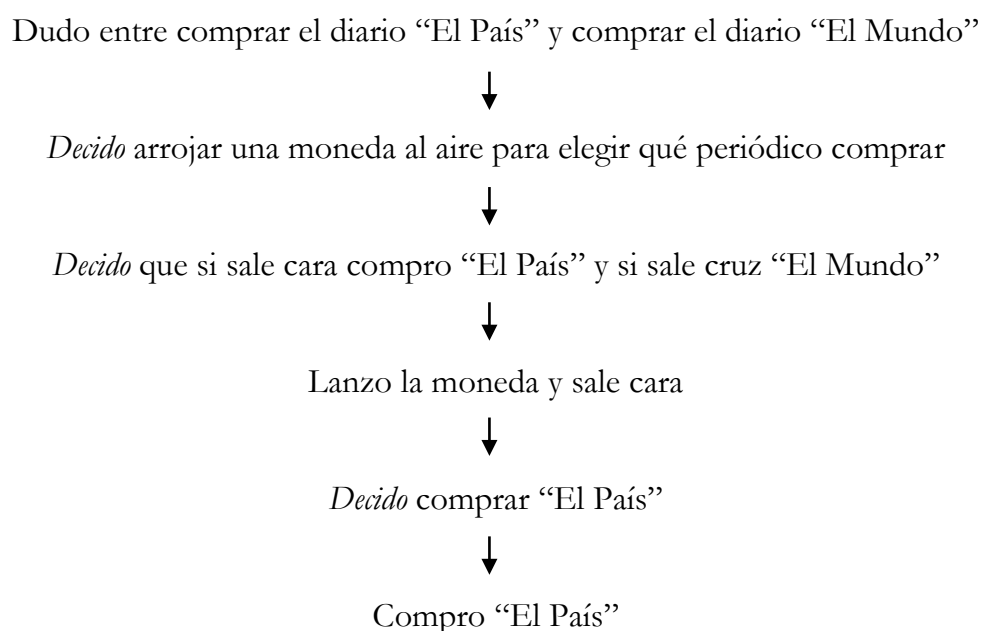
¹⁹⁶ Fischer 1999: 101.

Esta es tal vez una forma algo burda de presentar la postura libertarista, pero servirá de momento para ver cuál es el problema al que se enfrenta una concepción libertarista de la libertad, un problema no menos importante que el que tienen que resolver los compatibilistas. Resulta obvio que difícilmente puede hablarse de control del agente sobre sus acciones si éstas son meramente aleatorias. Supongamos que una persona decide arrojar una moneda al aire cada vez que tiene que elegir algo: ciertamente cada decisión suya motivada por el resultado de lanzar la moneda al aire no estaría determinada, pero tampoco sería el resultado de su control sobre la acción, sino precisamente de su falta de control sobre la misma. Normalmente arrojamus monedas al aire cuando no queremos responsabilizarnos por una elección, entonces dejamos “que la suerte decida”. Bien, en este ejemplo hay toda una serie de presupuestos no del todo correctos (tales como que el resultado del lanzamiento de la moneda no está determinado, la existencia del azar...), pero no corresponde ocuparse de ellos todavía, lo que interesa es señalar que no siempre que se habla de posibilidades alternativas (cara o cruz son dos posibilidades alternativas) se está cumpliendo correctamente con el *requisito* de las posibilidades alternativas pues éste entraña un control del agente sobre la acción (el segundo supuesto visto en la p. 94) que se pierde en las situaciones en que esas posibilidades vienen dadas por una indeterminación que es fruto del azar: “Precisamente, permitiendo las posibilidades alternativas en este sentido, uno atenúa la exigencia de que es el agente el que elige un sendero hacia el futuro.”¹⁹⁷

Naturalmente hay formas de indeterminismo mucho más sofisticadas que la que aquí he esbozado. La mayor parte de los libertaristas no aceptarían la idea de que el resultado de tomar una decisión arrojando una moneda al aire sea una elección libre: elegimos libremente arrojar la moneda, elegimos

¹⁹⁷ Ibidem.

libremente respetar su resultado, pero no elegimos ese resultado. La indeterminación debe hallarse en el lugar correcto de nuestro razonamiento práctico para ser relevante desde el punto de vista de la libertad. Por ejemplo, una secuencia de acción provocada por el lanzamiento de una moneda al aire podría ser la siguiente:



Según este esquema el que salga cara en la moneda no es mi decisión. Una vez ha salido cara cuento con un motivo más para comprar “El País”, pero siempre puedo pensar “¡Qué demonios, lo siento por la moneda, en realidad me apetece más leer «El Mundo»!”. Lo único que he hecho al decidir arrojar una moneda al aire es asumir que el resultado que depare dicha acción de arrojar esa moneda deberá ser *relevante* para mi decisión, puede que incluso decida arrojar la moneda con la sincera intención de que su resultado sea *determinante*; pero a menudo arrojamos monedas con la simple intención de que nos refuercen un deseo que prevalece frente a otro pero no lo suficiente como para empujarnos inmediatamente a la acción, o lo hacemos para descargarnos

ilusoriamente de responsabilidad y actuar con más tranquilidad.¹⁹⁸ En ocasiones sucede lo mismo cuando le pedimos la opinión a algún amigo sobre algo que queremos comprar. No es raro que lo único que queramos sea que su opinión coincida con la nuestra para comprar más a gusto, y que si su opinión no es la esperada la ignoremos y compremos lo mismo que ya teníamos pensado comprar pero enfurruñados (y no digamos ya si, además de su opinión, nuestro amigo nos da buenas razones para no comprar lo que ya teníamos pensado comprar).

Así, el libertarismo no consiste en postular sin más la inclusión del azar en nuestras vidas. Arrojar una moneda al aire supone (supuestamente) introducir una fuente de indeterminación en los *motivos* que cuentan para la elección y no en la elección misma. Pero una teoría libertarista sofisticada buscará un tipo de indeterminación que se dé en la elección misma, y no en los inputs que provoquen, tras una deliberación, el output que es la acción. Robert Kane, uno de los defensores actuales del libertarismo más destacados, lo explica así:

Hay una razón por la que [la indeterminación] debe estar entre el input y el output y no formar parte meramente del input. La razón es que todo cuanto suponemos que ocurre entre el input y el output corresponde a la actividad del agente (en la forma de razonamiento práctico que tendrá como resultado la elección). El input (en la forma de disposiciones, creencias, deseos y cosas por el estilo) no es algo que el agente controle aquí y ahora, aunque parte de ello pueda ser el producto de razonamientos, actividades o elecciones realizados en momentos anteriores [...]. El indeterminismo meramente en el input no da lugar a una responsabilidad definitiva.¹⁹⁹

En cualquier caso las teorías incompatibilistas (más o menos sofisticadas) siempre participan, en mayor o menor grado, de la idea de que la

¹⁹⁸ Se empieza a entrever aquí que en el incompatibilismo a menudo se mezclan dos planos, el epistémico y el ontológico, y que tal vez las tesis incompatibilistas sólo funcionan si estos planos se confunden. Que esta estrategia sea correcta es muy discutible hasta el punto, creo, de que demostrar su inconsistencia sería desarmar por completo al incompatibilismo. Ahora, a pesar de que aparezca aquí voy a continuar como si no tuviera esta intuición, pues será desarrollada detalladamente en I.3.

¹⁹⁹ Citado en Dennett 2004: 136.

libertad descansa en la posibilidad de acometer acciones sin determinación...²⁰⁰ lo cual puede entenderse no sólo como que el agente que lleva a cabo una determinada acción no está determinado, sino como que él mismo no es capaz de determinar su propia acción o de lo contrario no podría hablarse de posibilidades alternativas, y esto, como explica Leibniz en sus *Nouveaux Essais sur l'Entendement Humain*, es problemático: “Así la determinación nos es útil e incluso muy a menudo necesaria; y si estuviéramos poco determinados en toda suerte de encuentros y como insensibles a las razones (...) estaríamos sin elección efectiva (...).”²⁰¹ Generalmente tendemos a pensar, puesto que nada surge de la nada, que cuando el agente causa una determinada acción, entonces esa causa que es el agente es a su vez efecto de una causa anterior. Contra esta concepción se rebela el libertarista: “(...) los agentes debemos tener un tipo excepcional de libertad ‘metafísica’ o ‘contra-causal’. [...] debemos estar libres del nexo causal que determina otros sucesos naturales; debemos tener un tipo de libertad contra-causal que nos capacite para iniciar cadenas causales.”²⁰² Construir una noción de causa de este tipo es el gran reto del libertarista.

I.1.3. Sobre la causalidad no determinística

Como se acaba de ver, el requisito de las posibilidades alternativas nos enfrenta a dos graves problemas: por un lado parece incompatible con una causalidad de tipo determinista; pero por otro requiere una teoría de la acción basada en el poder causal del agente sobre sus actos, difícilmente asumible desde el indeterminismo. Así, para poder elaborar una teoría de la libertad de la voluntad que cumpla el requisito de las posibilidades alternativas, lo cual

²⁰⁰ Véase la nota 198. Seguramente no es lo mismo “sin determinación” que “sin conocimiento cierto de la determinación”. Insisto en que esta discusión ha de ser pospuesta hasta el capítulo I.3.

²⁰¹ Leibniz, *Op. Cit.* pf. 48, en Leibniz 1959: 263.

²⁰² Fischer & Ravizza 1993: 10.

considero indispensable, deberíamos ser capaces de llevar a cabo *al menos* una de estas dos cosas:

- O bien dar con una noción de causa compatible con el indeterminismo.
- O bien encontrar una noción de posibilidad alternativa compatible con el determinismo.

Aquí voy a optar por la segunda solución (de la que me ocuparé en *I.2.2.*), pero antes expondré algunas razones por las cuales no me parece viable, o cuando menos fácilmente asumible, la primera opción. Así, en este apartado *I.1.3.* voy a prestarle atención a las principales teorías libertaristas defendidas hoy en día, y en el próximo (*I.1.4.*) intentaré mostrar que, aún siendo muy iluminadoras acerca de muchos aspectos ocultos de la libertad de la voluntad, no resultan plenamente satisfactorias (a menos que uno esté dispuesto a asumir, y quedó dicho en la presentación que no es ese el caso, algún tipo de hipótesis metafísica fuerte sobre la existencia de sustancias diferentes de la materia y con otro tipo de propiedades).

Hecho este pequeño inciso, hay que señalar que las preguntas a las que debe responder aquel que decida acometer el proyecto de dar con una noción de causa compatible con el indeterminismo son las siguientes: “¿De dónde procede un acción si no es efecto de una cadena causal?”, “¿Qué es una causa que no determina?” o “¿De dónde surge una causa que no es determinada?”. Aunque la primera y la tercera se parecen, no son la misma pregunta, pues en la primera el libertarista está presuponiendo que toda acción es no causada, y en la tercera que toda acción es causada pero que el antecedente de cada acción no tiene a su vez antecedente. A cada una de estas preguntas le

corresponde un tipo de teoría libertarista, en las páginas siguientes voy a estudiar de forma pormenorizada las versiones actuales más importantes de los tres grandes modelos de teoría libertarista²⁰³, a saber:

- El *indeterminismo sencillo* de Carl Ginet.
- La *teoría probabilística* o del *indeterminismo causal*, tanto en la versión sencilla de Van Inwagen, como en las más sofisticadas de Robert Nozick y Robert Kane.
- La *teoría del agente causante* o de la *causación agencial* de Timothy O'Connor y de Randolph Clarke.

Aunque voy a hacer una caracterización del “indeterminismo sencillo” (*simple indeterminism*) que lo trata como una variante de la “teoría del agente causante” (*agent causal*), he optado sin embargo por mantener las citadas tres categorías porque tienen ya cierta tradición y recogen las tres vías por las que se ha intentado históricamente elaborar una teoría libertarista de la libertad de la voluntad. Así, considero que el “indeterminismo sencillo” es una forma de agustinismo, el “indeterminismo causal” una suerte de epicureísmo y la “teoría del agente causante” una construcción kantiana. Más adelante (en I.1.4.) justificaré esta categorización e intentaré mostrar que ninguna de estas tres vías consigue resolver el problema del control del agente. De momento voy a hacer un análisis crítico de estas teorías.

El indeterminismo sencillo

Se considera que el indeterminismo sencillo “mantiene que la libre agencia no requiere que haya ningún tipo de conexión causal (incluso de una

²⁰³ Tomo estas denominaciones de O'Connor.

variante indeterminista) entre el agente y sus acciones libres”²⁰⁴. El defensor del indeterminismo sencillo más citado y padre de la teoría es Carl Ginet; y la forma habitual de resumir su propuesta es como acabamos de ver, lo cual no deja de ser un punto simplista. Pero si uno atiende a las siguientes palabras de Ginet se da cuenta de que su teoría no es tan sencilla: “¿Son las explicaciones mediante razones (...) explicaciones *causales*? Lo son si todo lo que uno quiere decir con «causales» es que la explicación puede ser expresada mediante un «porque» uniendo la explicación y el explanans. [...] Si, por otra parte, uno necesita que una causa, propiamente dicha, preceda a su efecto, entonces no son explicaciones causales.”²⁰⁵

Así pues, el indeterminismo de Carl Ginet no es tanto un tipo de libertarismo sin causación, como uno con un tipo de causación que cabe resumir muy concisamente como “no-davidsoniano”.²⁰⁶ Precisamente la postura de Ginet es defender que el argumento del control del agente, que hace incompatibles el indeterminismo y la libertad, es incorrecto porque se basa en una definición incorrecta de “control” fruto de una idea de causación demasiado estrecha. Según Ginet las críticas al incompatibilismo son el resultado de “un caso de sobregeneralización de un paradigma bien entendido y muy respetado, en este caso el paradigma explicativo de las ciencias naturales en el que las leyes de la naturaleza son las que hacen conexiones explicativas. La fascinación hacia este paradigma puede, parece, cegarle a uno ante el hecho de que el paradigma explicativo de nuestras explicaciones corrientes de la

²⁰⁴ O'Connor 1995: 7.

²⁰⁵ Ginet 1989, en O'Connor 1995: 84.

²⁰⁶ A medida que vaya avanzando el trabajo habrá ocasión de profundizar en la teoría de la acción de Davidson, por el momento me limitaré a este breve resumen que de dicha teoría hace Carlos Moya: “En [«Acciones, razones y causas» Davidson] defendía la tesis según la cual las explicaciones de una acción mediante razones constituyen una forma de explicación causal, siendo las razones causas de la acción. Por otra parte, e inspirándose a este respecto en Elizabeth Anscombe, Davidson tendía a caracterizar implícitamente una acción intencional como aquella que tiene una explicación verdadera en términos de razones. Tal explicación ha de cumplir dos requisitos: las razones deben justificar racionalmente la acción y, además, deben causarla. Con ello, Davidson sentaba las bases de una concepción causal de la acción intencional humana según la cual una acción intencional es un proceso causal de cierto tipo, y se distingue de otros procesos por el tipo de causas que dan lugar a ella. Una acción intencional es un fragmento de

acción mediante razones es bastante diferente. Hay una relación interna y referencial que es suficiente para hacer la conexión explicativa y no es necesaria una conexión nómica. [...] Las explicaciones mediante razones no son *indeterminísticas*, sólo *adeterminísticas* (...).”²⁰⁷ A la vista de afirmaciones como esta última no es de extrañar que muchos autores consideren que Ginet habla de libertad sin causación. Aquí voy a defender, aun cuando esto vaya en contra de la interpretación canónica del indeterminismo sencillo, que puede decirse de la teoría de Ginet que en ella no hay conexión causal entre el agente y sus acciones si nos ceñimos a una noción davidsoniana o similar de causa, y que más bien la postura de Ginet lo único que hace (y por ello es tan original dentro del libertarismo) es que no admite que la causación indeterminística solucione el problema de la libertad. Frente a otros teóricos del libertarismo que, como se verá en breve, se afanan en demostrar que para que una acción sea libre ésta debe ser fruto de un yo que no es a su vez efecto de una causa diferente de sí mismo, o que ha sido causado indeterminísticamente; Ginet defiende el incompatibilismo no abogando por un tipo de causación indeterminista, sino sosteniendo que el patrón de las leyes naturales no es aplicable al funcionamiento agencial del ser humano y que por lo tanto, y he aquí su diferencia con Davidson, las razones de la acción poseen un estatuto causal propio diferente del de otros sucesos. En este sentido puede decirse que las razones de la acción no son causas; ir más allá es, considero, simplificar en exceso el indeterminismo sencillo. Ginet defiende que para la acción libre basta con una explicación racional que no tenga en cuenta causas anteriores a dicha acción, esto es, que no vayan más allá de la intención del agente de llevar a cabo esa acción: “(...) yo hago que mis propios, libres y sencillos [*simple*] actos mentales ocurran, no causándolos, sino simplemente siendo su sujeto, siendo

conducta cuyas causas son razones, en virtud de las cuales resulta justificado.” En Moya, C. “Introducción” a Davidson 1992: 14-15.

²⁰⁷ Ibidem: 91.

ellos mis actos. Son *ipso facto* determinados o controlados por mí, son suministrados libres, esto es, no determinados por alguna otra cosa, no causalmente necesarios por la acción de estados y sucesos anteriores.”²⁰⁸ Así pues, no es correcto apelar a motivos del agente previos a la acción para explicar la misma, dichos motivos sólo se dejan ver en el momento exacto de la acción. En cierto sentido motivos y acción, causa y suceso, son uno en lo que se refiere a la acción libre. De este modo Ginet consigue evitar el problema del control del agente: la acción se explica agencialmente por la determinación directa del agente, pero sin el concurso de ningún tipo de causa anterior. Nada más alejado de la noción de explicación causal de Davidson según la cual las razones primarias que impulsan la acción son un suceso diferente de la propia acción. Ginet estaría esgrimiendo el siguiente argumento, del que da cuenta Davidson en el artículo que le dio a conocer: “Puesto que una razón hace inteligible una acción al redescribirla, no tenemos aquí dos sucesos, sino uno en diferentes descripciones. Sin embargo, las relaciones causales exigen sucesos distintos.”²⁰⁹ Pero según Davidson las razones para actuar no son meras disposiciones sino causas empíricas, de lo contrario una racionalización de la acción, al ser una mera “redescripción”, vendría a no ser más que un enunciado analítico que sólo tendría sentido lógico.

Para eludir esta crítica Ginet recurre a la idea de intención: “Lo único *necesario* para la verdad de una explicación mediante razones (...), al margen del acontecer de la acción explicada, es que la acción haya venido *acompañada* por una intención con el tipo de contenido correcto.” De este modo Ginet evita caer en la autoreferencia de la explicación racional no-causal de la acción que denuncia Davidson, pero al mismo tiempo logra situar la causa de la acción en

²⁰⁸ Ginet, 1997: 89. Nótese que Ginet habla de actos no causados pero determinados. Creo que esto no hace sino reforzar mi interpretación del indeterminismo sencillo.

²⁰⁹ Davidson 1963, en Davidson 1995: 29-30.

el momento mismo de suceder la acción, esto es, consigue evitar que la acción se convierta en un efecto en el sentido de un suceso que cuenta con una causa antecedente. Bien, esos son los logros de la teoría de Ginet que parecen darle la razón a cierto tipo de incompatibilista, sin embargo esta aparente victoria deja, como vamos a ver, una serie de cadáveres por el camino que no son fáciles de asumir. En efecto, no es tan sencillo zafarse de la teoría de la agencia de Davidson a menos que uno cuente con un par de buenas respuestas que dar a las siguientes preguntas: ¿Cómo puede una causa no ser antecedente de un efecto y cómo es posible que causa y efecto se den simultáneamente?

Para Ginet el problema de la libertad no es con la causación, sino con la necesidad, con un tipo concreto de causa, y “¿si la conexión explicatoria entre la intención explicativa y la acción explicada no es necesidad nómica entonces que es? (...) [L]a intención concomitante explica la acción simplemente en virtud de ser una intención de esa acción por la que o en la que el agente hará una determinada cosa, en virtud, esto es, de ser ese tipo de actitud proposicional (una intención) cuyo contenido tiene ese rasgo [...] una relación *interna* entre el factor explicativo y la acción explicada. [...] La conexión explicatoria tiene lugar, no mediante leyes naturales, sino mediante la referencia directa y la relación interna.”²¹⁰ Así pues, lo que el indeterminismo sencillo defiende no es que actuar libremente sea actuar sin causa alguna, sino actuar por una causa que no entraña necesidad porque no es anterior sino simultánea a la acción misma de la que es causa. “Esto significa que tenemos un factor, el concomitante pretender algo de la acción en el agente, que es suficiente para verificar una explicación mediante razones de la acción y que no sólo no lo es sino que tampoco podría ser anterior a la acción.”²¹¹ En resumen: cada acción se explica a sí misma en virtud de sí misma.

²¹⁰ Ginet 1989, en O'Connor 1995: 83-84.

²¹¹ *Ibidem*: 82.

¿En qué se basa esta afirmación? Dado que lo que está defendiendo Ginet es que la relación que hay entre lo que hace que se dé la acción y la acción misma es de carácter *interno*, cabría suponer que se trata de una experiencia interna. La siguiente cita parece confirmar esta conjetura: “Toda acción, creo yo, o bien es o bien comienza con una acción mental causalmente sencilla [*simple*], esto es, un suceso mental que no consiste en un suceso mental causando otros. Un suceso mental sencillo es una acción si y sólo si tiene cierta *cualidad fenoménica intrínseca*, que he bautizado como la cualidad «actística» e intentado describir usando el vocabulario de la causación agencial adaptado radicalmente mediante «como si»: el suceso mental sencillo de mi volición de ejercer una fuerza con una parte de mi cuerpo fenoménicamente me parece intrínsecamente un suceso que no sólo me sucede, que no ocurre sin un mandato, sino que es, más bien, como si yo hiciese que ocurriera, como si yo determinase que vaya a suceder justo cuando y cómo lo hace(...).”²¹² En efecto, la idea de una “cualidad fenoménica intrínseca” sugiere que según Ginet lo que hace que una determinada acción, y no otra, pueda considerarse libre es que *se siente* como libre, uno experimenta la sensación de estar actuando sin otro motivo que la pura volición de llevar a cabo la acción. Se puede pues decir que, en cierto sentido, desde el punto de vista del indeterminismo sencillo la respuesta a una de las preguntas clave del libertarismo, “¿De dónde procede un acción si no es efecto de una cadena causal?”, es: de un motivo percibido por el agente como producido por él mismo y no determinado por ningún estado anterior. La clave está pues en la autopercepción: “En tanto me parece que he llevado a cabo ese suceso (en cierto sentido) directamente (...) no dudaría en suponer que lo he determinado.”²¹³

²¹² Ginet 1997: 89. La cursiva es mía.

²¹³ O'Connor 1993: 505.

Contra esta idea se pueden dar, al menos, dos buenos argumentos, o más bien un argumento (que extraigo de la tradición) y un buen ejemplo (que saco del debate actual) que lo corrobora. Podría llevarse la cuestión al terreno mucho más amplio del debate entre el internismo y el externismo, o continuar con los argumentos que a favor de una teoría davidsoniana de la agencia se pueden esgrimir (y de hecho un poco más abajo retomaré este punto). No obstante ello no parece necesario, pues el problema de la agencia interesa aquí de forma tangencial. Puede que la teoría de Ginet acerca de la acción intencional esté equivocada, pero no es necesario rebatirla, tan sólo estudiar si esa teoría es compatible con el control del agente y las posibilidades alternativas. Y creo que no lo es. Por ello bastará para rebatir el indeterminismo sencillo con unas palabras de Schopenhauer y un experimento mental ideado por Harry Frankfurt, que viene a reforzar el argumento del filósofo alemán. Ambos completan la siguiente sospecha que expresa otro gran teórico del indeterminismo, Timothy O'Connor, sobre la estrategia de Ginet: "(...) reconocemos que juicios de ese tipo elaborados sobre la base de la forma en que nuestro comportamiento es experimentado son, en principio, revocables. Pues se basan, en parte, en la presuposición por nuestra parte de que, en lo que al propio comportamiento se refiere, las cosas son básicamente lo que parecen, y por lo tanto que su *parecer* como si yo directamente llevase a cabo mis propias voliciones (...) se debe al hecho de que yo *de hecho* las llevo a cabo."²¹⁴ En efecto, no está tan claro que mi autopercepción como agente libre sea prueba suficiente de mi libertad de acción, y en cualquier caso no es sencillo aceptar sin más nuestra autoconciencia como garantía de libertad. Y precisamente con este problema que nos ocupa tiene mucho que ver la pregunta de la Real Sociedad Noruega de Ciencias a la que intentó responder Schopenhauer en su *Sobre la libertad de la voluntad*: "Nos dirigimos, pues, con

²¹⁴ Ibidem.

nuestro problema a la autoconciencia inmediata (...). ¿Qué luz nos arroja esta autoconciencia sobre aquella cuestión abstracta, a saber, sobre la aplicabilidad o no aplicabilidad del concepto de *necesidad* a la aparición del acto de voluntad luego de darse el motivo, esto es, de presentarse al intelecto?”²¹⁵

Resulta sorprendente hasta qué punto Schopenhauer coincide (si obviamos lo de “sin remedio”) con Ginet en señalar como contenido de la, siguiendo la terminología de Ginet, «cualidad actística», la siguiente actitud proposicional: “«Yo puedo querer, y cuando quiera una acción, los miembros movibles de mi cuerpo la ejecutarán en seguida, tan pronto como quiera, sin remedio». Más brevemente: *Puedo hacer lo que quiero*. Nada más aporta el testimonio de la autoconciencia (...).”²¹⁶ Aquello en lo que Schopenhauer no llegaría a estar de acuerdo con Ginet es en que este testimonio de la autoconciencia constituya una solución al problema de la libertad, pues para Schopenhauer, como se puede ver en la introducción del trabajo, el problema de la libertad no es una cuestión de “poder hacer” sino de “querer hacer”, y respecto a esta cuestión el testimonio de la autoconciencia no ofrece ninguna solución. Se puede entender que precisamente para evitar este problema Ginet intenta no hablar de causas, al menos en cuanto causas antecedentes: de ese modo una pista sobre el “poder” se convierte en la clave del “querer”. Desde el punto de vista del indeterminismo sencillo no hay realmente diferencia entre el querer, el poder y el hacer, los tres se dan simultáneamente. ¿Constituye esto realmente una solución al problema, o una ilegítima forma de postular la ausencia de antecedente de la voluntad, tan sólo para evitar caer en el problema de la regresión sobre un antecedente y otro y otro que acabaría llevándonos a postular el determinismo? “(...) [L]a autoconciencia declara la libertad del hacer, bajo el supuesto del *querer*, y la libertad por la cual preguntamos es precisamente la del *querer*. Nuestra investigación recae sobre la

²¹⁵ Schopenhauer 2000: 59.

²¹⁶ Ibidem: 60.

relación del querer mismo con el motivo, y sobre esta relación nada nos dice la afirmación «puedo hacer aquello que quiero». La dependencia de nuestro hacer, esto es, de nuestras acciones corporales, con respecto a la voluntad, que es lo que nos testimonia la autoconciencia, es algo muy distinto de la independencia eventual de nuestros actos de voluntad con respecto a las circunstancias exteriores, que es en lo que consistiría la libertad de la voluntad, y sobre esto nada nos puede decir la autoconciencia (...).”²¹⁷ En fin, lo que pienso, siguiendo a Schopenhauer, es que no es en absoluto irrelevante conocer los antecedentes de la acción para juzgar acerca de su libertad, no basta con el “estado mental sencillo” que tiene lugar al llevarse a cabo la acción. Ginet pretende no retroceder más allá de la acción misma para debatir acerca de su libertad, para hablar de posibilidades alternativas y de control del agente.

Creo, con Schopenhauer, que ello es insuficiente para comprender qué es actuar libremente, y una buena prueba a favor de esta intuición la suministra Harry Frankfurt en su “Alternate Possibilities and Moral Responsibility” donde propone el siguiente experimento mental: “Supongamos que alguien – digamos que Black- quiere que Jones ejecute cierta acción. [...] Así que espera hasta que Jones esté a punto de decidir qué hacer, y no hace nada hasta que tenga claro (Black es un juez excelente en este tipo de cosas) que Jones va a decidir hacer algo *diferente* de lo que él quiere que haga. (...) [En ese caso] Black toma las medidas necesarias para asegurarse de que Jones decide hacer, y de hecho hace, lo que él quiere que haga.”²¹⁸ ¿Cómo lograría Black llevar a cabo sus propósitos? Frankfurt nos propone que supongamos que Black es capaz de manipular el cerebro de Jones, de forma que una intervención en su sistema nervioso determine la acción de Jones. Nuestras intuiciones más inmediatas acerca de esta situación parece que nos obligan a decir que Jones no sería

²¹⁷ Ibidem.

²¹⁸ Frankfurt 1969, en Frankfurt 1988 : 6.

responsable de una acción que llevase a cabo por la intervención de Black. Pero Frankfurt complica las cosas: “Supongamos ahora que Black no tiene necesidad de dejarse notar porque Jones, por sus propias razones, decide realizar y realiza la acción que Black quiere que realice.”²¹⁹ Aquí ya no parece tan fácil decir si Jones es o no libre, las condiciones de su acción son las mismas que antes, pero Jones actúa “por sus propias razones”. El problema del indeterminismo sencillo es que obvia ese actuar “por sus propias razones”, dado que pone la marca de la libertad en el momento mismo del realizarse la acción, pues es en ese momento en el que se siente que lo que hace mi cuerpo es “como si” yo lo hiciese, y no como si me sucediera. Ahora, en el caso de Jones, ¿hay alguna diferencia en su autopercepción en el momento de actuar tanto libre como determinado por Black? Parece que ninguna, lo que distingue ambas situaciones es otra cosa. Jones, por efecto de la acción de Black en su cerebro, siente la acción que ejecuta pero que, de no haber mediado Black, no habría ejecutado, tan suya como si Black no hubiese existido. De hecho Black, para Jones, no existe. Lo que nos permite distinguir entre Jones haciendo A porque quiere y Jones haciendo A porque Black quiere no tiene nada que ver con el testimonio de la autoconciencia, porque lo que hace Black es que Jones *quiera* A y lo haga. En ambos casos Jones tendría la misma “cualidad actística”, la misma experiencia interna.

Así, es necesario recurrir a los antecedentes de una acción para decidir acerca de su libertad, y esto atenta contra el núcleo del indeterminismo sencillo, cuya originalidad reside precisamente en fijarse en la decisión sencilla, singular, aislada, que provoca la acción. Parece que la determinación causal forma parte de la explicación de una acción precisamente en tanto entendemos la causa como antecedente, y esa es precisamente la propiedad del concepto de causa que Ginet no está dispuesto a admitir. Esta dificultad que tiene Ginet

²¹⁹ Frankfurt 1969, en Frankfurt 1988 : 7.

para aceptar algún tipo de determinación causal de las acciones es fruto, como se vio más arriba, de su temor a la idea de que sea posible formular leyes naturales que expliquen y predigan las acciones humanas. Dicho temor es el que hace que muchos autores, desde los filósofos helenistas, hayan optado por hablar de determinación sin necesidad. No obstante la teoría de la acción de Davidson contempla esa posibilidad y constituye una respuesta coherente y fiable a esos temores. ¿Qué necesidad hay de negar el carácter causal de la agencia si dicho carácter causal no entraña la posibilidad de formular leyes científicas al respecto? Criticar la teoría davidsoniana por el peligro de una naturalización excesiva de la acción humana es precisamente no entender la originalidad de la solución de Davidson, que “(...) admite (...) que no hay conjunciones constantes o leyes estrictas que conecten las razones con las acciones. Lo que hay, a lo sumo, son generalizaciones vagas y llenas de excepciones. Pero señala que esto mismo sucede en el caso de muchos enunciados singulares de carácter claramente causal.”²²⁰ El que exista una relación causal entre dos sucesos es independiente de que exista una ley que asuma esa relación y la generalice, lo primero tiene que ver con hechos, lo segundo con nuestros conocimientos acerca de los hechos. “La ignorancia de leyes predictivas aptas no inhibe la explicación causal válida; de lo contrario podrían hacerse pocas explicaciones causales. Estoy seguro de que el cristal de la ventana se rompió porque una piedra lo golpeó, yo vi cómo sucedió todo; pero no dispongo (y ¿quién dispone?) de las leyes sobre cuya base pueda predecir qué golpes romperán qué ventanas.”²²¹

He querido reseñar este aspecto porque más adelante defenderé que a menudo se quiere sacar unas consecuencias del materialismo y del

²²⁰ Moya, C. “Introducción” a Davidson 1992: 18.

²²¹ Davidson 1963, en Davidson 1995: 32.

determinismo que van mucho más allá de lo necesario,²²² y la teoría de la agencia de Davidson es una buena muestra de lo erróneo de ese punto de vista. No obstante considero que el indeterminismo sencillo es incorrecto al margen de estas consideraciones; básicamente su problema reside en que se basa en los datos inmediatos de la conciencia “[y] nada nos puede decir la autoconciencia sobre la regulada necesidad que hace que de dos deseos contrarios uno de ellos, y no el otro, se convierta en acto de voluntad y en acción, ya que esa autoconciencia se entera del resultado *a posteriori*, sin que lo sepa *a priori*.”²²³

El indeterminismo causal

Puede decirse que esta forma de libertarismo es la más intuitiva, o cuando menos la más común. A la pregunta “¿Qué es una causa que no determina?”, se responde “una causa indeterminística”. Cuando Van Inwagen piensa en una teoría incompatibilista de la libertad de la voluntad, piensa sobre todo en una teoría probabilística de este tipo. Con matices, los partidarios del indeterminismo causal, según O’Connor, vienen a defender lo siguiente: “(...) el agente controla sus acciones mediante sus razones previas. Tales razones (...) debe pensarse que causan, pero no implican necesariamente, la acción. [...] [E]l que tales acciones sean un producto de «quienes somos» [no requiere] que las acciones sean un resultado *determinístico* de estados mentales previos, tan sólo un resultado causal.”²²⁴ O dicho de otro modo: “un agente causa sus acciones libres via sus razones para actuar de ese modo, pero indeterminísticamente”.²²⁵ De aquí resulta que, lo que para la mayor parte de los teóricos del

²²² Generalmente ello es debido a que se mezclan dos cuestiones, a saber: la de cuáles son los hechos y la de cuáles pueden ser mis conocimientos acerca de los hechos. En *I.3.1.* trataré de deshacer ese error.

²²³ Schopenhauer 2000: 61.

²²⁴ O’Connor 1993: 507-508.

indeterminismo causal vendría a sustentar empíricamente la idea de causas no determinantes, sería la teoría cuántica.²²⁶ Ésta permitiría introducir cierto grado de indeterminación en los procesos causales de toma de decisiones, pero con la ventaja de que a escala macro esa indeterminación no podría leerse como intervención del azar en la susodicha toma de decisiones. Así pues, esta perspectiva está basada en la idea de que, a indeterminaciones cuánticas, pueden sobrevenir variaciones en el comportamiento previsible de objetos macroscópicos. De este modo sería posible decir que siempre existe alguna probabilidad de que alguien no actúe como está previsto que actúe, a pesar de la acción de las leyes naturales sobre sus antecedentes. Algunos autores hablan de un indeterminismo que debe ser leído como “casi determinismo”. Basta con que sea posible ocasionalmente la libertad, aunque sea de forma excepcional. En cualquier caso esta caracterización del indeterminismo viola menos palmariamente el requisito del control del agente que versiones del mismo que, como mostré en *I.1.2.*, acaban equiparando la libertad al azar. Además, los partidarios del indeterminismo causal suelen esforzarse por distinguir cuáles son las ocasiones en que realmente está algo en juego cuando se trata de la acción libre, y esto es interesante porque dice mucho acerca de cómo se constituye un sujeto a sí mismo por sus acciones, y cómo, una vez constituido éste, sus acciones vienen en gran medida dadas por aquello en que el sujeto se ha convertido. De este modo, esta estrategia incompatibilista conseguiría dar cuenta del requisito del control del agente además de cumplir, por definición, el requisito de las posibilidades alternativas. Un requisito éste que, desde el indeterminismo causal, no habría de verse cumplido siempre para que nos fuera posible hablar de libertad, tan sólo en los momentos apropiados.

²²⁵ O'Connor 1995: 7.

²²⁶ Como he anticipado en la nota 222, más adelante, en *I.3.1.*, pondré en duda la idea de que la indeterminación se da en los hechos cuánticos y no en nuestro conocimiento. De momento me ceñiré a la lectura que de la mecánica cuántica hacen los incompatibilistas interesados en una noción indeterminista de causalidad.

En esta línea Van Inwagen, y esta idea la comparten todos los libertaristas que vamos a estudiar en este apartado, considera que no todas nuestras acciones son genuinamente libres. Según Van Inwagen generalmente lo que elegimos hacer depende de la información que poseamos, de nuestro deseo de hacerlo y de que lo consideremos factible física y moralmente (que no veamos ningún impedimento para la acción en cuestión). La mayor parte de nuestras acciones vendría, pues, determinada por estas circunstancias y no podría ser considerada genuinamente libre. Otras muchas acciones serían realizadas sin deliberación, y tampoco serían apropiadas para hablar de libertad. “Ahora, hay *algunas* ocasiones en las que un agente se enfrenta a alternativas y no tiene claro qué hacer (...).”²²⁷ Estas ocasiones pertenecen a tres categorías. La primera de ellas es irrelevante, se trata de casos como el del asno de Buridán que, como sabemos, era incapaz de saciar su hambre ante dos montones de paja, equidistantes e igualmente apetecibles, decidiéndose por uno de ellos. Esta categoría está compuesta por situaciones que otros incompatibilistas (verbigracia, Carl Ginet) han tendido a considerar como aquellas en que, dado que en ellas las decisiones se toman desde una indeferencia total, la libertad de la voluntad aparece más claramente. Ejemplos típicos son: decidir coger el teléfono con la mano izquierda y no con la derecha, o al revés; y escribir un asterisco en el ordenador usando la tecla de encima del teclado numérico en lugar de la tecla +, o al revés. Van Inwagen considera que dichas situaciones reflejan efectivamente elecciones entre posibilidades alternativas no determinadas, pero que son moralmente irrelevantes, y por ello las deja de lado. Otra categoría de elecciones genuinamente libres la componen aquellas en que se da una “lucha moral”, y otra más las indecisiones de carácter existencial (lo que toda la vida se han

²²⁷ Van Inwagen 1989, en O'Connor 1995: 233.

llamado “dilemas”). Los casos de “lucha moral” son aquellos en que no somos capaces de decidirnos entre un deseo que podría reportarnos múltiples satisfacciones de forma inmediata, y lo que consideramos nuestro deber. Las indecisiones de carácter existencial son aquellas decisiones en que está en juego qué queremos ser y es imposible decidirse porque el resultado de cualquiera de las elecciones posibles es deseable por sí mismo, pero no es comparable con el resultado alternativo. Un ejemplo clásico de este tipo de situaciones es el que pone Sartre en *El existencialismo es un humanismo* sobre un joven francés que, en plena Segunda Guerra Mundial, está dividido entre quedarse cuidando a su inconsolable madre, abandonada por su marido y con un hijo muerto en la guerra, o enrolarse en las Fuerzas Francesas Libres en Inglaterra. Así, “(...) hay al menos *dos* tipos de ocasiones en que el incompatibilista puede admitir que ejercemos la libertad de la voluntad: casos de una lucha real entre lo percibido como deber moral o interesante para uno a largo plazo, de un lado, y un deseo inmediato, del otro; y casos de conflicto entre valores inconmensurables.”²²⁸ ¿Y cómo se ejerce en esos casos dicha libertad de la voluntad? Según Van Inwagen la libertad se manifestaría en casos como el que sigue. A un ladrón le sucede que, en ocasiones, cuando está robando, le vienen a la mente su madre agonizante y la promesa que él le hizo en su lecho de muerte de que iba a ser una persona honesta. En esas ocasiones el ladrón decide no robar, y según Van Inwagen su decisión es indeterminada, esto es, “hay mundos posibles en los que las cosas podrían ser absolutamente idénticas en cada aspecto a la forma en que son en el mundo real en el momento en que nuestro arrepentido ladrón toma su decisión –mundos en los que, es más, las leyes naturales son exactamente tal cual son en el mundo real– y en los que toma el dinero”.²²⁹ La idea de Van Inwagen es que en algunas situaciones de la vida hay tan buenas razones para hacer A como para hacer B,

²²⁸ Ibidem: 235.

²²⁹ Van Inwagen 1983: 128.

y que en esas ocasiones no está determinado qué hará el agente, o cuando menos puede decirse que tan posible es que suceda A como que suceda B.

A propósito de este tipo de fórmulas que equiparan posibilidades dentro de la deliberación Richard Double ofrece, en “Libertarianism and Rationality”, un buen contra-argumento (que dirige no sólo contra Van Inwagen, sino también contra Kane).²³⁰ Double opina que las teorías libertaristas, en aras del indeterminismo, no salvaguardan la racionalidad de las elecciones libres. Así, Double considera que hay un principio, el “Principio de la Explicación Racional” (PRE), que es tan débil que incluso los libertaristas deberían admitirlo. El principio es el que sigue: “Referirse al proceso de razonamiento R de una persona explica racionalmente una elección C sólo si la probabilidad de C dado R es mayor que la probabilidad de no C dado R.”²³¹ En efecto, si lo racional es actuar conforme a las razones que nos permiten elegir entre cursos de acción alternativos, entonces, si ciertas razones R_A conllevarían normalmente el que hiciésemos A y no B, decir que hay tantas posibilidades de que ocurra A como de que ocurra B, dado R_A , es contradecir el hecho mismo de que R_A conlleve normalmente A y no B, pero esa es la definición misma de R_A . Pues bien, según este modelo tan básico de explicación, no podría decirse que las decisiones tomadas desde la teoría de la decisión de Van Inwagen fueran racionales. Según Van Inwagen el ladrón está dividido entre su deseo de robar y su deseo de cumplir la promesa que le hizo a su madre; y si cupiese hacer una media de todas las ocasiones en que roba y en que decide no robar, entonces hallaríamos que hay tantas posibilidades de que el ladrón robe como de que no robe.

Así que parece que el argumento de más abajo es aplicable directamente a Van Inwagen:

²³⁰ He optado por exponer este argumento aquí y no al final de este apartado porque Kane ha refinado su teoría después de que el argumento de Double fuera presentado, con lo que habría que revisar dicho argumento antes de incluirlo aquí como un falsador de la teoría de Kane.

²³¹ Double 1988, en O'Connor 1995: 60.

- (i) Si tenemos una libertad de la voluntad libertaria, entonces nuestras deliberaciones no dotan a nuestras decisiones más que de un 0,5 de probabilidad.
- (ii) Si el PRE es verdadero, entonces si nuestras deliberaciones no dotan a nuestras decisiones más que de un 0,5 de probabilidad entonces nuestras deliberaciones no explican nuestras decisiones.
- (iii) El PRE es verdadero.
- (iv) Así pues, si tenemos una libertad de la voluntad libertaria, entonces nuestras deliberaciones no explican nuestras decisiones.²³²

Ante críticas así Van Inwagen acaba admitiendo que defiende el modelo probabilista más por reducción al absurdo que por considerarlo plenamente correcto: “Es por ello que mi título es «El *misterio* de la libertad metafísica». [...] Mi opinión es que el primer argumento (el argumento por la incompatibilidad de la libertad y el determinismo) es básicamente correcto y que hay, por tanto, algo mal en el segundo argumento (el argumento por la incompatibilidad de la libertad y el indeterminismo). Pero si me preguntan *qué* es, tengo que decir que estoy, como se dice en el argot, totalmente despistado.”²³³ Van Inwagen considera que el argumento de la consecuencia es plenamente correcto y decide apostar por el modelo probabilista: “(...) ciertamente no es inaudito en filosofía que un argumento incontrovertible le empuje a uno a una conclusión enigmática para la que uno no tiene explicación (...). He de elegir entre lo enigmático y lo inconcebible. Elijo lo enigmático.”²³⁴ En cualquier caso, dados los tipos de situación a los que Van Inwagen restringe la libertad, da la impresión de que, mientras que su parte negativa (la del argumento de la consecuencia) está bien trabada, su parte positiva es más bien fruto de un compromiso moral (el mismo que seguramente tenemos todos exceptuando al escéptico moral, suponiendo que éste exista), que resultado de argumentos convincentes. Por ello no puede tomarse la teoría probabilística de Van Inwagen como modelo del indeterminismo causal, creo que este autor es más

²³² Ibidem: 61.

²³³ Van Inwagen 2002, en Kane 2002: 194.

²³⁴ Van Inwagen 1983: 150.

interesante como crítico del compatibilismo que como defensor del libertarismo.

Pero existen versiones más sofisticadas del indeterminismo causal, construcciones positivas más interesantes que las meras buenas intenciones de Van Inwagen. Una de ellas es la de Robert Nozick, que está íntimamente ligada a su bien conocida teoría de la decisión. Según Nozick, cuando uno ha de decidir hacer algo, sopesa ciertas razones que se corresponden con ciertas acciones posibles y no con otras; llegado el momento de actuar unas razones acaban por pesar más que otras y provocan su acción correspondiente: “(...) las razones no vienen con pesos especificados con precisión dados previamente; el proceso de decisión no es de descubrimiento de pesos precisos de ese tipo sino de asignación de los mismos. El proceso no sólo sopesa razones, (también) las carga. Al menos, así parece a veces. Este proceso de carga se concentraría estrechamente en, o implicaría considerar o decidir, qué tipo de persona desea uno ser, qué tipo de vida desea uno llevar.”²³⁵ La idea de Nozick es que el acaecer de la deliberación se corresponde con los procesos descritos por la mecánica cuántica. Así lo explica Timothy O'Connor:

Nozick (...) sugiere que la interpretación ortodoxa de la teoría de la medición de la mecánica cuántica suministra un modelo análogo estructuralmente a este análisis de la decisión como adjudicadora de pesos comparativos a razones no clasificadas previamente. En esa teoría, un sistema mecánico cuántico está en una «superposición» continuamente cambiante, o una mezcla probabilística, de estados, que cambia discontinuamente en el momento de una medición. Se dice de la medición que «desbarata el paquete de ondulaciones», reduciendo la superposición a un estado concreto de forma indeterminada. De manera análoga, en la teoría de Nozick, una persona tiene razones con pesos no especificados previamente a la decisión (así puede decirse que está en una superposición de pesos con respecto a sus razones). El proceso de decisión reduce la superposición a al menos una clasificación comparativa determinada, y no está determinado a qué clasificación de ese tipo será reducida.²³⁶

Por lo tanto se puede hablar de un modelo indeterminista. El problema es que si el peso de las razones es indeterminado, no causado por principios

²³⁵ Nozick 1981, en O'Connor 1995: 102.

²³⁶ O'Connor 1993: 511.

motivadores previos del agente, entonces la decisión no parece ser fruto del tipo de control del agente que es relevante para la libertad, sino fortuita: “(...) incluso aunque la teoría del indeterminismo causal (...) admita la posibilidad real de cursos de acción diferentes, dentro de los cuales cualquiera pueda ser «controlado» por el agente en el sentido de que pueda ser el producto causal de las razones que pertenecen al agente, no parece ‘depender del agente’, algo ‘sobre lo que tiene elección’, el determinar exactamente *qué* causa potencial será eficaz para cada instancia dada, y por tanto qué acción tendrá lugar realmente.”²³⁷ Nozick es consciente de este problema: “El que una acción sea no-determinada no es suficiente para que ésta sea libre –podría ser simplemente un acto aleatorio. Si actuásemos como el uranio 238 emite partículas alpha, el determinismo sería falso pero (a menos que estemos totalmente equivocados acerca del uranio 238) con eso no tendríamos libertad de la voluntad.”²³⁸ Por ello Nozick introduce en su teoría un principio de decisión: cada uno escoge lo que cree que es mejor. El problema es que este principio, si está ya antes de la decisión, supone una forma de evaluar ya presente en el sujeto que dota a priori de determinado peso a cada razón en base a lo mejor o peor que sea. Dicho de otro modo, ese peso, conocido o no de antemano por el agente, ya existiría, las razones serían a priori mejores o peores. Pero lo que hacía interesante la propuesta de Nozick, al menos para los libertaristas, era que hasta el momento de la decisión las razones no adquirirían su peso, es más, las razones no tenían peso, sólo el agente se lo daba: decidir sería cargar después de sopesar. Por ello Nozick defiende que el principio de decisión es a su vez creado en cada elección, es *natura naturans* y *naturata*: la política de hacer lo mejor confiere “pesos a razones sobre la base de la decisión tomada en el momento de elegir una concepción de uno mismo y de la vida que es apropiada para uno, una concepción que incluye el conferir

²³⁷ Ibidem: 508.

²³⁸ Nozick 1981, en O'Connor: 105.

esos pesos y el elegir esa concepción (en que los pesos también se conforman a la elección de esa auto-concepción).”²³⁹ De este modo Nozick conseguiría que su recurso a la mecánica cuántica, necesario para salvar las posibilidades alternativas, no conduzca a una teoría en que las decisiones fueran fruto del azar. Sin embargo cabe dudar de que a la pregunta “¿Por qué hizo X?” se pueda responder de forma satisfactoria con “Porque X es una acción adecuada a los principios que están presupuestos en el hacer precisamente X”. No se me ocurre otra forma hacer comprensible esta respuesta que convirtiéndola en un argumento ontológico *à la* Descartes, estrategia ésta que dudo suscriba Nozick y que sin embargo tiene un cierto aire de familia con la suya: “Supongamos que tenemos una decisión en que los pesos que se confiere «fijan principios generales que ordenan no sólo el acto que es relevante sino también la asignación de esos pesos (o similares)». ¿Podemos explicar cómo depende del agente el que esos pesos (antes que otros) sean asignados, señalando que la decisión es «una instancia de la propia concepción y pesos escogidos»?”²⁴⁰ Si el principio de decisión de cada uno es la ley que nos permite sopesar y cargar las razones que nos mueven a la acción, entonces no puede instituirse por la decisión misma que se supone es resultado de su existencia previa. Así explica Sartre este punto en *El Ser y la Nada*:

Resulta de ello que la deliberación voluntaria siempre está falseada. En efecto: ¿cómo apreciar motivos y móviles a los cuales precisamente yo confiero su valor antes de toda deliberación y por la elección que hago de mí mismo? [...] En realidad, los móviles y los motivos no tienen sino el peso que les confiere mi proyecto, es decir, la libre producción del fin y del acto conocido por realizar. Cuando delibero, los dados ya están echados. Y, si debo llegar a deliberar, es simplemente porque entra en mi proyecto originario darme cuenta de los móviles *por medio de la deliberación* más bien que por tal o cual otra forma de descubrimiento (...).²⁴¹

El principio de decisión, *qua* principio, explica cómo funciona la acción, y no puede ser al mismo tiempo el porqué de una determinada acción y su

²³⁹ Ibidem: 106.

²⁴⁰ O'Connor 1993: 513.

²⁴¹ Sartre 1943: 476.

resultado; es una norma general del “cómo”, no una causa particular de cada uno de los “por qué”, que a su vez lo convertirían en causa “universal”. Si el principio de decisión funciona como explicación de *la* acción, no puede servir como determinante de *esta* acción, como dador de peso que inclina la balanza hacia una acción concreta en vez de otra: se trata de una base del *actuar*, no de *una acción* concreta y más tarde de otra y de otra y de otra. Si esto fuese así habría tantos principios de decisión como acciones, puesto que cada acción, además de basarse en él, *funda* el principio de decisión. Existe una “desanalogía entre una ley explicativa, y una decisión que *instituye* una política general o una concepción de acuerdo con las cuales uno debería actuar. Las leyes explicativas no se hacen verdaderas en un momento dado. De haber una analogía aquí, sería entre decisiones que se subsumen a sí mismas y un suceso de una ley haciéndose verdadera, suceso que sería explicado por, puesto que estaría subsumido por, el contenido real de la ley.”²⁴² Aclararé este confuso trabalenguas con otro: lo que es correcto no puede convertirse en correcto por el hecho de ser adoptado como correcto en un momento dado, sobre la base de que es principio de lo correcto. Así pues la teoría de Nozick, siendo más sofisticada que la de Van Inwagen, tampoco logra ser plenamente satisfactoria, pues no consigue dar cuenta del requisito del control del agente, en tanto su explicación de cómo pueden depender de uno las decisiones que se toman, siendo éstas fruto de un sistema de deliberación que no incluye clasificación alguna que determine la elección, tiene la forma de una petición de principio. Por ello puede decirse que la teoría de Nozick no consigue aclarar cómo podría considerarse que el agente tiene algún control sobre sus acciones. Hay que admitir, pues, que tampoco esta versión del libertarismo consigue eludir el argumento según el cual indeterminismo y libertad, en tanto ésta requiere de control del agente, son incompatibles.

²⁴² O'Connor 1993: 513.

Queda por observar, dentro del indeterminismo causal, una de las teorías libertaristas más influyentes de los últimos años, la de Robert Kane. Para este filósofo la solución al problema de la libertad está también en la mecánica cuántica, que permitiría introducir indeterminación en nuestras deliberaciones sin que dicha indeterminación se notase a nivel macro, esto es, en nuestras explicaciones racionales de la acción. Así, Robert Kane es partidario del indeterminismo pero con cortafuegos. Además Kane admite, con Van Inwagen, que la libertad es algo raro que no acaece siempre, sino tan sólo en situaciones en que tenemos dudas acerca de qué acción llevar a cabo. Eso sí, Kane reduce esas situaciones a las de indecisión existencial, a los dilemas, y deja de lado los casos de conflicto entre deberes y deseos inmediatos. Sólo en caso de dilema tendríamos realmente posibilidades alternativas porque la capacidad de “obrar de otro modo” implicaría, según Kane, un poder de “autocontrol racional dual”, esto es, “poder para elegir cualquiera de los dos caminos como el resultado razonable de un proceso deliberativo”²⁴³, y únicamente en situaciones en que las posibilidades que se nos ofrecen son igualmente razonables, como en los dilemas, puede darse ese hecho. En cierto sentido se trata de situaciones en que más que elegir qué queremos hacer, tenemos que elegir quién queremos ser (y vuelve a resonar Sartre por aquí).²⁴⁴ En “Two Kinds of Incompatibilism”²⁴⁵ Kane señala tres condiciones indispensables para que una teoría libertaria dé cuenta adecuadamente de procesos decisorios y fundamente la responsabilidad:

(RU) Un agente tiene *responsabilidad última* por la elección de A o alguna alternativa (esto es, depende “del agente aquí y ahora” en el sentido de responsabilidad última el que el agente elija A o elija de otro modo), sólo en caso de que:

²⁴³ Kane 1985: 60.

²⁴⁴ Un poco más adelante desarrollaré esta parte de la teoría de Kane con más atención.

²⁴⁵ Recojo aquí la reformulación que hace O'Connor de estas condiciones porque la encuentro más clara que la del propio Kane.

- (1) (Condición de la producción) La elección que se haga sea el fin intencional de un esfuerzo de la voluntad que sea el esfuerzo de la voluntad del agente.
- (2) (Condición de la racionalidad) El agente (r_1) tiene razones para elegir así, (r_2) lo hace por *esas* razones, (r_3) no elige (por esas razones) de manera compulsiva, y (r_4) en el momento de la elección cree que las razones por las que ésta se hace son en algún sentido las de más peso, que merece la pena actuar por ellas más que por sus alternativas.
- (3) (Condición de la exclusividad) Dados los hechos de la situación, ninguna otra explicación (aparte de la conjunción de (1) y (2)) es posible para la elección del agente, a menos que esa explicación pueda pasar a ser explicada ella misma por la conjunción de (1) y (2). (I.e., la explicación suministrada por (1) y (2) es «exclusiva» o «final».) En concreto, cualquier explicación de cómo el agente hace un esfuerzo de la voluntad en (1) y de cómo el agente tiene el carácter y las razones o motivos para elegir en (2) tampoco explicaría la elección, aun cuando (1) y (2) explicaran la elección.²⁴⁶

Tanto la condición (1) como la (2) están pensadas para evitar el problema del control del agente, ya que, desde estas condiciones el desarrollo de una elección parece ser un proceso racional de decisión que controla el agente y que no tiene lugar aleatoriamente. La condición (1) además, y ahora se verá con calma por qué, tiene por objeto que esta teoría de la decisión no se vea falsada por el argumento de la consecuencia. Así, puede decirse que especialmente la condición (2) es una condición “determinista”, y la (1) “indeterminista”. La condición (3) es meramente procedimental, no hace sino poner límites a la proliferación de condiciones, al mismo tiempo que defiende que (1) y (2) son igualmente necesarias para explicar la acción (recuérdese, por cierto, que estoy hablando tan sólo de aquellos procesos de decisión en que todas las alternativas son igualmente razonables). En fin, (1) introduce en la elección el grado de indeterminación necesario para que haya posibilidades alternativas, y (2) describe el nexo existente entre la elección y las razones del agente para que exista control por parte de éste sobre sus acciones. A primera vista tenemos la teoría perfecta.²⁴⁷ Veamos ahora hasta qué punto es esto cierto estudiando en qué consiste el “esfuerzo de la voluntad”.

²⁴⁶ O'Connor 1993: 517. El subrayado es mío.

²⁴⁷ Y desde mi punto de vista efectivamente la de Kane es la teoría libertarista más sólida, y por ello le voy a dedicar una parte importante de este trabajo a su teoría.

Algo producto del esfuerzo de la voluntad es algo que sucede porque intentamos que suceda, algo que tiene lugar porque nos ponemos en disposición de que ocurra. Por ejemplo: una buena idea para una novela puede ocurrirnos casualmente mientras estamos cocinando, o mediante el *esfuerzo de la voluntad* ante una página en blanco. En el primer caso se puede decir que la idea simplemente acaece, en el segundo es provocada por nuestro buscar y rebuscar una buena idea hasta dar con ella. Tal vez un símil tenístico pueda ayudar a entender mejor este punto de la teoría de Kane: la diferencia que hay entre una acción que es fruto del esfuerzo de mi voluntad y una que no lo es, es la misma que hay entre ganar un punto por un fallo del contrario provocado por un pericioso golpe mío, y ganarlo por lo que se suele llamar un “error no forzado”. “(...) [Los] «esfuerzos de la voluntad» en la terminología de Kane no son «voliciones» en sentido tradicional, sino «esfuerzos mentales dirigidos a conseguir que los fines de uno (propósitos, intenciones) sean escogidos».”²⁴⁸ Ahora, ¿por qué el “esfuerzo de la voluntad” beneficia al libertarista? Porque interviene en el proceso de decisión, lo causa, pero es indeterminado:

El grado de esfuerzo es realmente indeterminado. Con todo, si es suficiente, entonces hago suficiente esfuerzo, y si no es suficiente, entonces no hago suficiente esfuerzo. Puedo golpearme por no haber hecho más (...). No puedo decir que hice cierta (determinada) cantidad de esfuerzo y entonces fue cosa de suerte si esa cantidad de esfuerzo fue suficiente. Uno no puede decir eso porque el esfuerzo realizado fue él mismo indeterminado. De forma similar, uno no podría decir de dos personas en exactamente la misma situación, que hicieron la misma cantidad de esfuerzo y, por azar o suerte, uno de ellos logró no caer en la tentación mientras que el otro no lo hizo. [...] El esfuerzo y la suerte no pueden ser separados y vistos como variables independientes. La suerte es simplemente el ser indeterminado del esfuerzo. Es por esto que no es posible absolverse de responsabilidad en situaciones de elección moral basándose en la mala suerte, aunque el resultado no esté determinado.²⁴⁹

Así, en la teoría de Kane cada vez que nos encontramos ante una situación en que hay posibilidades alternativas, la libertad de nuestra elección depende del esfuerzo de la voluntad que hagamos. ¿Un esfuerzo por qué? Un esfuerzo por que se nos ocurran posibilidades alternativas que vayamos a

²⁴⁸ O'Connor 1993: 518.

²⁴⁹ Kane 1985: 154.

preferir a aquellas que, de momento, aparecen en nuestra mente. “Es una especie de *esfuerzo de la voluntad* «dejar las puertas de la conciencia abiertas a nuevas o ya consideradas posibilidades que pudieran surgir del inconsciente, de la memoria, o de la libre asociación de ideas, y pudieran orientarnos en la deliberación práctica o en un problema [creativo] por resolver». Ese esfuerzo de la voluntad sería él mismo indeterminístico.”²⁵⁰ Pongamos por ejemplo que al ver un accidentado en la carretera se me ocurren dos posibilidades: parar mi coche y atenderle, o continuar como si nada hubiera pasado.²⁵¹ Pero resulta que existe una tercera posibilidad, tal vez la mejor, que no se me ocurre: llamar con mi móvil a una ambulancia. No se me ocurre esta posibilidad porque soy dueño de un teléfono móvil desde hace poco tiempo y no estoy acostumbrado a recurrir a él, pero el caso es que llevo el móvil guardado en la guantera. La idea de Kane es que yo soy responsable de esforzarme por que se me ocurra la posibilidad de usar el móvil para pedir ayuda, pero el que se me ocurra esa posibilidad, me esfuerce yo mucho o poco, *es algo indeterminado*. Sin embargo no por ello soy menos responsable: hay una clara diferencia, siguiendo con el ejemplo, entre que yo me olvide del accidentado pero a los pocos kilómetros alguien me llame por teléfono, suene el móvil, y entonces caiga en que puedo usarlo para pedir ayuda, y que yo dé con la solución del móvil como resultado del seguir dándole vueltas a las posibles soluciones a la situación a la que me enfrento, esto es, como resultado del esfuerzo de mi voluntad. De este modo la condición (1) cumple con el control del agente, porque lo único que hace es que sea posible que se me ocurra una forma de actuar que yo juzgaré correcta, y que por lo tanto será fruto de mis deseos y creencias, de mis razones para actuar. Pero además (1) cumple con los anhelos del libertarista de que ese control no implique determinación. Esto es ir mucho más allá de las formas

²⁵⁰ O'Connor 1993: 516.

más o menos vagas que habíamos visto hasta ahora de hablar de causas que no determinan:

[N]o basta con decir que mi decisión dependió de mí porque fue causada, pero no causada necesariamente, por (una o más de) mis razones. Pues debería darse el hecho de que dependiese de mí –al menos en parte– el que esa razón causara mi decisión en ese caso concreto. Se nos revela entonces que este requerimiento se cumple en virtud del hecho de que el funcionar esa relación no-necesitante en ese momento fue influido por un esfuerzo de la voluntad. En cuanto al esfuerzo de la voluntad propiamente dicho, bien, fue *mi* esfuerzo, ¿no?²⁵²

El esfuerzo de la voluntad es una causa *mía* de *mi* acción, y es indeterminada. Así pues, ya están cumplidos todos los requisitos para la libertad de la voluntad. “Chapó” para Kane. ¿O no? Cabe un par de dudas acerca del esfuerzo de la voluntad como parte relevante de la deliberación.

En primer lugar uno puede preguntarse en qué beneficia finalmente el citado esfuerzo a la libertad. Pongamos que estamos dando vueltas y más vueltas con el coche por el centro de Madrid intentando aparcar, tarea ésta realmente ardua y enervante. Llega un momento en el que nos planteamos dejar el coche en el parking, pero la perspectiva de que cena, cine y parking sumarán demasiados euros nos hace esforzarnos y seguir con nuestra búsqueda. Por fin un coche deja su sitio libre y aparcamos. ¿Quiere esto decir que si hubiésemos aparcado antes de llegar a dudar entre dejar el coche en el parking o continuar, el haber aparcado habría sido una acción menos libre? Y sobre todo, si la indeterminación viene dada por el hecho de que no está determinado por nosotros qué coche, cuándo y dónde, va a dejar su sitio, ¿en qué beneficia la indeterminación a la libertad? ¿Qué cambio supone para mi acción? Nosotros somos nuestro coche, no el centro de Madrid. ¿En qué sentido es más libre mi acción si lo que se me ocurre hacer es indeterminado? Hago el mismo esfuerzo de la voluntad cuando decido salir de casa para ir al

²⁵¹ Desde luego aquí no hay dilema moral alguno, una posibilidad es deseable y la otra no. El objetivo de este ejemplo es tan sólo ilustrar en qué consiste el esfuerzo de la voluntad y por qué según Kane éste es relevante para una teoría de la decisión.

²⁵² O'Connor 1993: 521.

centro que cuando venzo la tentación de entrar en el parking. No es ahí, ni en el coche que me deja su sitio, donde queda la libertad. Tal vez el esfuerzo de la voluntad forme parte a menudo de nuestras decisiones, incluso en aquellas ocasiones en que Kane considera que tiene lugar verdaderamente la libre elección; pero está por ver que nuestra libertad sea fruto *precisamente* de ese esfuerzo, o más bien que coincida con ese esfuerzo, con la ocurrencia de ideas sobre cursos de acción posibles, una ocurrencia de la que no somos causantes, a lo sumo facilitadores. Da la impresión de que la libertad, si existe en nuestras elecciones, está en otro lugar del proceso.

En segundo lugar, hay un problema de regresión al infinito²⁵³ al reducir, lo que tienen de libre nuestras elecciones al “esfuerzo de la voluntad”:

Por supuesto Kane puede señalar que el hecho de que ese esfuerzo de la voluntad tenga lugar no es simplemente un asunto de aleatoriedad –son el propio carácter del agente y motivos en conflicto los que inducen la «lucha» interna. Pero la forma concreta en que el esfuerzo de la voluntad acaba revelándose no está a su vez determinada por su carácter previo y motivos. Podría evolucionar de forma diferente. Así que si el agente es responsable de cómo ese periodo de deliberación cobra forma, de que suceda que evolucione por ese camino en vez de por algún otro camino que podría disminuir significativamente la probabilidad de que la decisión que se termine tomando triunfe, entonces debe depender de él (...) que esas circunstancias previas causen (aunque no causen necesariamente) que ese esfuerzo de la voluntad tenga ese carácter concreto más bien que otro. No podemos resolver adecuadamente este problema en un estado de actividad (la decisión de uno) simplemente refiriendo al que pregunta a otro anterior (un «esfuerzo de la voluntad»);²⁵⁴

Así, parece que Kane no ha hecho sino resituar el problema de la libertad pero, como acaba de verse con el ejemplo del aparcamiento imposible en Madrid, está por ver que esa nueva situación ayude realmente a resolver la cuestión. Para demostrar que sí, Kane ha sofisticado su teoría en una obra realmente admirable, *The Significance of Free Will*. Intentaré resumir lo más sumariamente posible los nuevos argumentos de Kane.

²⁵³ Recuérdese que en el caso de la teoría de Nozick sucedía lo mismo, parece que la circularidad es un rasgo definitorio del indeterminismo causal, probablemente porque dichas teorías parten del hecho que quieren demostrar, a saber, que la libertad se da dentro del indeterminismo.

²⁵⁴ O'Connor 1993: 521. He cambiado los tiempos verbales para que este párrafo sea comprensible fuera de su contexto.

Se ha visto más arriba que Kane habla de “responsabilidad última” del agente, y ha llegado el momento de tratar con más detenimiento esta idea. Puede decirse que en la teoría de Kane esta “responsabilidad última” hace las veces de concepto de control del agente con una diferencia: engloba, además del requisito del control de agente propiamente dicho, el requisito de las posibilidades alternativas. Así, según Kane no puede tenerse control sobre las acciones sin que ello conlleve posibilidades alternativas, y por ello Kane se centra en el problema del control del agente o, en sus términos, de la responsabilidad última. “La idea básica es ésta: para ser responsable último de una acción, un agente debe ser responsable de cualquier cosa que sea razón suficiente (condición, causa o motivo) de que ocurra la acción. Si, por ejemplo, una elección se sigue de, y puede ser explicada suficientemente por, el carácter y los motivos (junto con las condiciones de fondo) de un agente, entonces para ser responsable *último* de la elección, el agente debe ser al menos en parte responsable, en virtud de elecciones y acciones realizadas voluntariamente en el pasado, por tener el carácter y los motivos que tiene ahora.”²⁵⁵ Esto no hace sino recoger, según Kane, la idea de Aristóteles de que un hombre que sufre de debilidad de la voluntad, en algún momento debió formar él mismo el carácter débil del cual surge esa voluntad, y que por lo tanto dependía de él ser o no débil. Así, ser responsable último sería, para Kane, lo mismo que poseer libertad de la voluntad, siendo ésta “el poder de los agentes para ser los creadores y soportes *últimos* de sus propios fines y propósitos”.²⁵⁶ Por ello, y vuelvo ahora a otra cuestión que mencioné cuando empecé a estudiar la teoría de Kane, no es necesario que la libertad de la voluntad, con su contenido de posibilidades alternativas, presida todas nuestras acciones; tan sólo que esté presente en situaciones excepcionales en que está en juego que el agente se autoconstituya como tal, en situaciones en que se forma el carácter. Se trata de

²⁵⁵ Kane 2001b, en Kane 2001a: 224.

²⁵⁶ Ibidem: 224.

las “acciones auto-formadoras”. Uno podría estar determinado por su carácter para hacer algo en un momento concreto y bajo unas condiciones dadas, y aún así ser responsable de dicha acción, pero “sólo en la medida en que fuera responsable de sus motivos y carácter presentes en virtud de muchas luchas previas y acciones auto-formadoras (AAFs) que le llevaron a ese punto en que no puede hacer otra cosa. [...] A menudo actuamos desde una voluntad ya formada, pero es «nuestra propia libre voluntad» (...).”²⁵⁷ En fin, según Kane, para considerarnos responsables siempre y bajo cualquier circunstancia que no sea de coerción, de ignorancia, de enajenación mental... basta con que pueda decirse de nosotros que somos responsables de ser quienes somos. Sin acciones auto-formadoras en nuestra biografía es imposible que podamos ser considerados responsables. Una vez más resuena Sartre en esas palabras: “El para-sí es necesario en tanto que se funda a sí mismo.”²⁵⁸

Ahora bien, esas acciones auto-formadoras, ¿son determinadas por mí? Entonces no puedo ser libre en sentido libertarista. ¿Son indeterminadas? Entonces no puede decirse que yo tenga control alguno sobre quien quiero ser. Como se ha dicho poco más arriba, la idea de Kane es que para que pueda considerarse que somos seres libres basta con que las acciones auto-formadoras sean indeterminadas, basta con que tengamos posibilidades alternativas en momentos críticos de nuestra vida en que “estamos partidos entre visiones rivales de lo que deberíamos hacer o en lo que deberíamos convertirnos”.²⁵⁹ En los dilemas soportamos una gran tensión debida a una absoluta indecisión entre dos posibilidades igualmente deseables o indeseables, Kane sugiere que esa tensión “se refleja en regiones apropiadas de nuestros cerebros por un desvío del equilibrio termodinámico –brevemente, una suerte de «actividad caótica» en el cerebro que lo hace sensible a

²⁵⁷ Ibidem: 225.

²⁵⁸ Sartre 1943: 117.

²⁵⁹ Kane 2001b, en Kane 2001a: 228.

microindeterminaciones a nivel neuronal”.²⁶⁰ Así, cuando el Cid, en la obra de Corneille, está dividido entre vengar a su humillado padre enfrentándose a Don Gómez, o no hacerlo para no tener que renunciar al amor de Gimena, hija de Don Gómez; es porque esa misma división está teniendo lugar en el cerebro de Rodrigo, sus neuronas están sometidas a la misma incertidumbre, a la misma indeterminación, que su “alma”, que ha de elegir entre el hijo y el amante. “Lo que es experimentado internamente como incertidumbre corresponde así físicamente al abrirse de una ventana a la oportunidad de temporales bajadas del telón de la completa determinación por influencias del pasado. (Por contra, cuando actuamos desde motivos predominantes o disposiciones ya fijadas, entonces la incertidumbre o indeterminación está silenciada. Si jugase un papel en tales casos, *sería* una mera molestia o chiripa, tal y como sostienen los críticos del indeterminismo.)”²⁶¹ En el caso del Cid sus motivos para actuar, el amor por su padre y el amor por Gimena, están en conflicto, de tal forma que le es imposible actuar determinado por cualquiera de ellos:

Que je sens de rudes combats!
Contre mon propre honneur mon amour s'intéresse:
Il faut venger un père, et perdre une maîtresse:
L'un anime le coeur, l'autre retient mon bras.
Réduit au triste choix ou de trahir ma flamme,
Ou de vivre en infâme,
Des deux côtés mon mal est infini.²⁶²

En principio la opción más tentadora es la que atiende a Gimena, pero el deber de Rodrigo, por su honor, es ante todo vengar la ofensa que le han hecho a su padre. Así, Rodrigo necesita un *esfuerzo de la voluntad* para vencer la

²⁶⁰ Ibidem: 228.

²⁶¹ Kane 2001b, en Kane 2001a : 228.

²⁶² Corneille 1636 (301-307). Puede traducirse estos versos así: “¡Qué combates tan rudos siento! / Contra mi propio honor mi amor se interpone: / Hay que vengar a un padre, y perder a una señora: / Uno anima el corazón, la otra frena mi brazo. / Reducido a la triste elección o de traicionar mi llama, / O de vivir en la infamia, / De los dos lados mi mal es infinito.”

tentación de no enfrentarse al padre de Gimena. Si logra vencer la tentación será porque hizo ese esfuerzo, si no, porque no lo hizo, “y esto se debe al hecho de que, mientras que quería resistir a la tentación, también quería caer en ella, por razones bien diferentes e inconmensurables”.²⁶³ En este sentido la decisión del Cid está indeterminada, pero ello no quiere decir que sea aleatoria, “no se sigue, de que uno no controle o determine cuál, de un conjunto de resultados, vaya a ocurrir antes de que ocurra, que uno no controle o determine cuál de ellos ocurra, *cuando* ocurra. [...] [L]os agentes ejercen control sobre sus vidas futuras *en este momento y en este lugar* decidiendo. Realmente tienen (...) «control voluntario plural» sobre las opciones en el sentido siguiente: son capaces de realizar *cualquiera* de las opciones que quieran, *cuando* quieran hacerlo, por las razones que quieran, y a propósito más bien que accidentalmente o por error (...)”²⁶⁴ El indeterminismo que crea la incertidumbre se corresponde pues con el hecho de que haya varias opciones. Al haber varios deseos confrontados hay varias corrientes neuronales abiertas que luchan por llevarse el gato al agua de causar el output que es la acción. Pero esto no implica que el output sea indeterminado, será igualmente determinado por un motivo o por otro, y por ello es necesario que las redes neuronales que se corresponden con los motivos tengan una “actividad caótica”.

El problema es el siguiente: “La «actividad caótica» que menciona Kane es un caos *determinista*, la impredicibilidad *práctica* de cierta clase de fenómenos que pueden describirse perfectamente dentro de la vieja física newtoniana.”²⁶⁵ Esta actividad caótica entraña simplemente que, cuando hay demasiadas cosas que intervienen en la acción, ésta está sometida a tantas influencias diferentes que es imposible predecir el resultado. Un buen ejemplo es del un bombo de

²⁶³ Kane 2001b, en Kane 2001a: 228.

²⁶⁴ Ibidem: 230-231.

²⁶⁵ Dennett 2003: 127.

lotería: la posición final de una bola está determinada por su posición en el bombo antes de que éste empiece a rotar, y por todos los rebotes que sufra mientras el bombo esté en movimiento. Para que fuese posible calcular el movimiento de la bola, habría que calcular con absoluta precisión los movimientos de las bolas restantes, lo cual no es posible. Así, lo que produce el caos del movimiento de la bola es un exceso de influencias relevantes para su movimiento, un movimiento que no por impredecible deja de estar determinado. ¿Y por qué, entonces, habría de ser importante este caos? Porque permitiría explicar el hecho de que el Cid escogiera el amor de Gimena en vez de lo que es su deber, esto es, vengar a su padre. La actividad caótica de nuestras redes neuronales explicaría, por ejemplo, la debilidad de la voluntad; explicaría que aún queriendo hacer A, sabiéndolo nuestro deber, antes que B, hiciésemos B. ¿Pero en qué beneficia esto a la libertad? Supongamos que el Cid, meditabundo, desea vengar a su padre, pero coincide que Don Diego, padre de Rodrigo, ha cogido la capa de Don Gómez y la lleva puesta. Así, Rodrigo (supongamos que por una vez actúa sin honor) asesta una poderosa estocada por la espalda a alguien que lleva la capa de Don Gómez, pero que resulta ser Don Diego. En este caso también contamos con una explicación para el hecho de que el resultado de la acción de Rodrigo no se corresponda con sus deseos, pero no por ello consideramos que, gracias a ello, puede hablarse más claramente de libertad en su acción. Supongamos que siempre que el Cid hubiese de matar a alguien, no tuviese otra referencia que la capa de la víctima. Y añadamos como condición de “actividad caótica” que las víctimas potenciales estuviesen constantemente intercambiando sus capas. ¿Hacen estas condiciones más libre al Cid? Parece que no.

Por supuesto, lo que podría objetar Kane es que estas condiciones han de darse en el cerebro del Cid, no en el mundo. Y tal vez una objeción similar podría dirigirse contra el ejemplo que desarrollé más arriba de un desesperado conductor buscando dónde aparcar en Madrid. ¿Pero por qué el hecho de que

esas condiciones fueran internas habría de hacerlas más relevantes? Tanto si la incertidumbre viene de dentro, como si viene de fuera, no es más que una interferencia para la libertad. Lo que ocurre es que si las condiciones son internas, entonces parece que el esfuerzo de la voluntad es relevante para la acción, mientras que si son externas, no. Pero si es la ausencia de esfuerzo de la voluntad lo que explica que el Cid se quede con Gimena, y por el contrario el que haya hecho un esfuerzo de la voluntad suficiente es lo que explica que venga a su padre, ¿por qué es necesaria la actividad caótica? Lo que explicaría en cualquier caso que hagamos A y no B sería la ausencia o presencia de un esfuerzo de la voluntad. Pero entonces, como se dijo más arriba, lo único que estaría haciendo Kane sería resituar el problema: de acuerdo, la deliberación no es indeterminada porque ello haría que nuestras decisiones fuesen irracionales (según vimos más arriba, viendo el “Principio de la Explicación Racional”), lo es el esfuerzo de la voluntad. Pero resulta que el esfuerzo de la voluntad, o bien es parte de la deliberación, y entonces estamos en las mismas, o bien, si es externo a la deliberación, tendría el mismo efecto que el constante cambio de capas de la víctimas del Cid. Por ello no parece ni que la actividad caótica de nuestras redes neuronales, ni el esfuerzo de la voluntad, puedan ser la solución al problema del libertarismo, con lo que cabe concluir que el indeterminismo causal, tal y como ha sido desarrollado hasta ahora, no es una teoría que dé cuenta del requisito del control del agente, y por lo tanto no es una teoría válida acerca de la libertad de la voluntad.

El agente causante

Los partidarios de esta teoría tienen, con respecto a las acciones indeterminadas, la misma duda que William James atribuye a los partidarios del determinismo: “Si el acto «libre» significa una extraña novedad, que no procede *de* mí, el yo previo, sino *ex nihilo*, y simplemente se añade a mí, ¿cómo podría yo, este yo previo, ser responsable? ¿Cómo puedo yo tener *carácter*

alguno permanente capaz de persistir el tiempo suficiente para ser recompensado con alabanza o reprobación? El rosario de mis días se desharía en infinitud de cuentas tan pronto como el hilo de la necesidad interna fuera suprimido por la absurda doctrina indeterminista.”²⁶⁶

Probablemente los partidarios de la teoría del agente causante suprimirían lo de “absurda doctrina indeterminista”, pero por lo demás las palabras de James expresan bastante bien sus inquietudes. “(...) [P]arece que yo mismo libre y directamente controlo el resultado, y aquí (como en todas partes) «control» es evidentemente una noción *causal*. Y el hecho de que el indeterminismo causal no es satisfactorio sugiere que tenemos que ser bastante literales respecto al referente de «yo» en este contexto. Si hago algo libremente, no puedo ser visto simplemente como el lugar en el que factores internos y externos trabajan juntos para producir mi acción.”²⁶⁷ Así, la teoría del agente causante pretende, desde el punto de vista del incompatibilismo, salvar el escollo del control del agente convirtiendo a éste en el fundamento de todas nuestras acciones: “propone una forma *sui generis* de causación por un agente que es irreducible (tanto ontológica como conceptualmente) a procesos causa-efecto dentro del agente”.²⁶⁸ Algunos autores²⁶⁹ leen este principio como una apuesta por una teoría de la acción davidsoniana; el lector avezado caerá inmediatamente en que, dada la crítica que hice a Ginet al analizar el indeterminismo sencillo, no es esa mi impresión. La teoría de la acción de Davidson distingue entre causas y razones para actuar, en nuestra forma de percibirlas y describirlas, pero en ningún caso entre géneros diferentes de causación “tanto ontológica como conceptualmente”. Del mismo modo que en su teoría del monismo anómalo Davidson no defiende que haya dos

²⁶⁶ James, 1907: 83.

²⁶⁷ O'Connor, 2001: 199.

²⁶⁸ O'Connor, 1995: 7.

²⁶⁹ Véase Kane 1989. Una lectura de Davidson en la línea de la defendida en este trabajo puede verse en Ginet 2001.

realidades *sui géneris*, lo físico y lo mental, sino una sola realidad a la que en ocasiones nos es dado aplicar el lenguaje causal de la ciencia, y en ocasiones el lenguaje intencional de la conciencia, así, en su teoría de la acción tampoco está postulando dos tipos de causación esencialmente diferentes. Lo que hace Davidson con el monismo anómalo es poner límites al naturalismo, frenar el reduccionismo, pero no por ello abrazar hipótesis dualistas. El rechazo del eliminativismo, en la teoría de la mente y de la acción de Davidson, no puede ser leído como una apuesta por una distinción sustancial entre mente y cerebro o razones y causas. Tal vez pueda leerse que la teoría de Davidson defiende que la causación propia de las acciones es “irreducible conceptualmente” a procesos de causa-efecto, pero no que es “irreducible ontológicamente”. Así, igual que mi sospecha inicial sobre el indeterminismo sencillo se asentaba en la teoría davidsoniana de la acción, en el caso que nos ocupa ahora, por mucho que los defensores de la teoría del agente causante intenten escudarse en ella, se puede fundar una nueva sospecha sobre la misma base. Hecha esta aclaración puede proseguir el estudio de la teoría del agente causante.

Los partidarios de esta teoría tienen por padre fundador a Thomas Reid, y las primeras versiones recientes, si se me permite la expresión, de esta forma de incompatibilismo se deben a C. D. Broad, en un primer momento, y a Richard Taylor y Roderick Chisholm, después. De Broad se pueden extraer unos principios generales de la teoría del agente causante, más precisos que la definición básica que ya se ha visto: “Broad define la causación agencial, a la que llama «causación no-aconteciente», como la agencia de un yo con respecto a sus elecciones o acciones libres que satisfaga las siguientes condiciones. (a) El yo o agente es la única causa de sus elecciones o acciones libres; (b) su causación puede ser ejercida en dos direcciones, elegir (o hacer) y hacer otra cosa; y (c) su causación de una elección o acción libre es la causación de un acontecimiento o suceso por una *cosa* o *sustancia* que no puede ser explicada como la causación de un acontecimiento o suceso *por otros acontecimientos* o

sucesos.”²⁷⁰ Como bien dice Kane comentando estos principios, todos los libertaristas estarían de acuerdo con (a) y (b), así que lo realmente característico de la teoría del agente causante es (c). (b) es la parte de la teoría destinada a cumplir el requisito de las posibilidades alternativas y (c) la destinada a cumplir el del control del agente sobre sus acciones. Así, la respuesta a la pregunta “¿De dónde surge una causa que no es determinada?”, que caracteriza a la apuesta libertarista por la teoría del agente causante, vendría a ser: del agente. Esta es la letra de la teoría del agente causante, su espíritu aparece bien reflejado en esta muestra del pensamiento de Chisholm: “Si somos responsables, y si lo que hemos estado intentando decir [acerca de que no son lo mismo acciones y sucesos] es cierto, entonces tenemos una prerrogativa que algunos tan sólo atribuirían a Dios: cada uno de nosotros, cuando actúa, es un primer motor inmóvil. Haciendo lo que hacemos, causamos el que ciertos sucesos tengan lugar, y nada –o nadie– nos causa para que causemos que esos sucesos ocurran.”²⁷¹ Hay que señalar, en aras del rigor, que curiosamente tanto Taylor como Chisholm han abandonado recientemente sus teorías originarias,²⁷² pero en cualquier caso no iba a ocuparme aquí de ellas, sino de dos propuestas más recientes: la de Timothy O’Connor y la de Randolph Clarke. Dichas propuestas participan del espíritu del proyecto de Broad, Chisholm y Taylor, pero han limado sus defectos más visibles.

Hasta ahora hemos visto a O’Connor en su faceta de exegeta del libertarismo. En lo que respecta a su propia teoría puede considerarse que se trata de una visión de la causación agencial bastante tradicional. “Randolph Clarke ha puntualizado que la explicación tradicional de la causación agencial tiene dos rasgos sobresalientes. Primero, la causación agencial es una especie

²⁷⁰ Kane 1989, en O’Connor 1995: 117-118. Broad usa indistintamente “suceso” y “acontecimiento”.

²⁷¹ Chisholm 1964, en Watson 1982: 11.

genuina de causación, pero fundamentalmente diferente en su naturaleza de la causación sucesiva [*event causation*].²⁷³ Segundo, el suceso causado agencialmente no es pensado como sujeto también a causación sucesiva.”²⁷⁴ A estos rasgos generales hay que sumarles algunos trazos más para dar con la teoría completa de O’Connor. Según dicha teoría los objetos son dinámicos y poseen intrínsecamente propiedades causales, de tal forma que intervenir en procesos causales no es algo que meramente le suceda a los objetos, sino una tarea que éstos emprenden. Lo que hace que la causación agencial y la sucesiva sean diferentes es la forma en que esas capacidades son ejercidas. En el caso de la causación sucesiva “el *ejercicio* de una capacidad o una tendencia tal procede automáticamente: un objeto que tiene, en las circunstancias adecuadas, el racimo de propiedades que es su fundamento capacitador, genera directamente uno de los efectos que está a su alcance.”²⁷⁵ En lo que respecta a la causación agencial “(...) la capacidad también está circunscrita a factores físicos y psicológicos que operan mientras el agente delibera. Pero (y aquí está la diferencia con el paradigma mecanicista) tener las propiedades que ayudan a la capacidad de causación *agencial* no es suficiente para producir un efecto particular (...); más bien, *faculta* al agente para determinar un efecto (...). Si, cuándo y cómo ha de ser ejercida dicha capacidad es determinado libremente por el agente.”²⁷⁶ Y esas condiciones que determinan el momento y la forma en que ha de tener lugar la acción componen la intención del agente, además del hecho de que esa acción sea resultado de ciertas razones. Y aquí es donde empieza a tener problemas O’Connor, pues si la intención explica la acción, y la intención incluye tan sólo razones, entonces no consigue explicar el

²⁷² A este respecto se pueden consultar Chisholm 1995, y Taylor, R. 1982.

²⁷³ Adopto “causación sucesiva” como traducción de “event causation”. Sería más correcto hablar de “causación de sucesos”, y como tal debe ser comprendida la “causación sucesiva”, pero he preferido dotarme de una expresión antagónica de “causación agencial” que pueda funcionar gramaticalmente en castellano de modo análogo a como lo hace en inglés.

²⁷⁴ Fischer 1999: 107.

²⁷⁵ O’Connor 2001: 200-201.

siguiente hecho, muy acertadamente señalado por Davidson: “(...) alguien puede tener una razón para una acción y, sin embargo, realizar la acción sin que esta razón sea la razón por la que la hizo. La idea de que el agente realizó una acción *porque* tenía una razón es fundamental para la relación entre una razón y la acción que explica.”²⁷⁷ ¿De dónde puede O’Connor sacar ese “porque” que explique la acción, sin recurrir a la causación sucesiva? Pues bien, O’Connor recurre, como ya se vio hiciera Ginet aunque de forma diferente (pues se trataba de una intención concomitante, no causante), a la idea de intención. “En la perspectiva de la causación agencial de O’Connor, una acción consiste en que el agente causa el que una intención mueva su cuerpo, de alguna manera, *inmediatamente*, intención que causa y sustenta el movimiento del cuerpo.”²⁷⁸ Así, O’Connor conseguiría suministrar a su teoría un nexo causal que permitiría dar cuenta de la carencia que según Davidson caracteriza a las explicaciones de las acciones que no recurren a la causalidad. Pero el problema es entonces que O’Connor ya no puede concebir las causas que llevan a actuar al agente como meramente facultativas: “(...) es difícil ver cómo una perspectiva causalista de este tipo podría encuadrarse en la teoría de O’Connor según la cual un suceso que consista en un agente causando algo no podría a su vez ser causado.”²⁷⁹ En efecto, hemos visto que O’Connor defiende que, dadas las causas capaces de producir una acción, el agente está capacitado para decidir que dichas causas efectivamente produzcan la acción, y que por ello puede considerarse que el agente tiene control sobre sus acciones. Pero si acaba admitiendo, para superar el argumento de Davidson, que la acción se explica por su causa (la intención, compuesta por las razones del agente), y no por la decisión de que la causa opere, entonces O’Connor está perdiendo lo que le es valioso en su explicación: que el agente genera la acción

²⁷⁶ Ibidem: 201.

²⁷⁷ Davidson 1963, en Davidson 1995: 24.

²⁷⁸ Ginet 2001: 393.

ayudado por las causas de la misma, pero no *a causa* de ellas. El problema de la teoría de O'Connor reside pues en que el agente parece ser efecto al mismo tiempo que causa, pero esto contradice los postulados mismos de la teoría del agente causante. Así pues O'Connor debería, o bien renunciar a esos postulados, o bien encontrar un contraejemplo para el argumento de Davidson a favor de la utilidad del concepto de causa para explicar la acción. No parece dispuesto a hacer lo primero y no ha sido capaz de lograr lo segundo. Por tanto se debe concluir que la teoría de O'Connor está, de momento, en un callejón sin salida.

La teoría de Randolph Clarke difiere de las concepciones tradicionales de la teoría del agente causante y recibe el nombre de “perspectiva del causal agente-causante”. Precisamente su originalidad radica en que, ante el problema que acabamos de ver, Randolph Clarke reacciona renunciando a uno de los postulados de la teoría del agente causante: el de la diferencia intrínseca entre causación agencial y sucesiva. Según su teoría “la causación agencial implica exactamente el mismo tipo de causación que la causación sucesiva. La única diferencia reposa en los relata de la relación causal: en el caso de la causación sucesiva, ambos relata son sucesos, mientras que en el caso de la causación agencial, el primer relatum es un agente y el segundo un suceso.”²⁸⁰ Así, aquello por lo que es especial la causación agencial es por el hecho de que sus efectos son producidos por una sustancia, y no por las propiedades mismas de dicho tipo de causación. Este punto de vista permite a Clarke superar uno de los problemas de las teorías tradicionales del agente causante, que, cómo acabamos de ver en el caso de la de O'Connor, no consiguen dar cuenta del hecho de que mi actuar pueda explicarse por mis razones *porque es causado* por mis razones. En el fondo la originalidad de la propuesta de Clarke reposa en el

²⁷⁹ Ginet 2001: 393.

²⁸⁰ Fischer 1999: 107.

hecho de que su teoría *sí* es davidsoniana, a pesar de defender la importancia del agente causante. Sólo hay un tipo de causación, pero hay diferentes formas de explicarla porque no todas las causas son iguales: una sustancia bien puede ser una causa, pero tiene una serie de propiedades que harán que, probablemente, su forma de causar no pueda ser estudiada exactamente del mismo modo en que lo es la forma de causar de un simple suceso. La idea de Clarke es que el agente no es lo único que determina la acción, sino algo que la codetermina: “(...) [E]l que el agente actúe con libertad de la voluntad consiste (crucialmente pero no sólo) en que su acción sea causada, en este sentido, por él y por sus razones.”²⁸¹ Bien, de este modo Clarke consigue evitar los peligros de las teorías tradicionales del agente causante... lo que ya no está tan claro es que sea capaz de escapar al argumento de la consecuencia. La versión libertarista de Clarke acaba estando tan cerca del compatibilismo como del incompatibilismo. En efecto, mientras que los incompatibilistas tienen problemas con el control del agente, Clarke parecería poder llegar a tenerlos más bien, como les ocurre a los compatibilistas, con las posibilidades alternativas pues: ¿qué sentido tendría afirmar que una misma causa, un mismo agente, una misma sustancia, con unas mismas razones, habría podido causar una cosa y su contraria? Por ello Clarke termina defendiendo que, cuando el agente actúa libremente, es porque su llegar a tener las razones que le llevan a actuar causa indeterminísticamente su acción. Dado que la causalidad agencial es una causalidad sucesiva, entonces está regida por leyes naturales, pero no por leyes de cualquier tipo: “(...) aquel tipo concreto de acción que el agente realiza libremente estaría gobernado por leyes causales, pero por leyes causales indeterminísticas más bien que determinísticas.”²⁸² La libertad reposaría en el hecho de que el agente decide actuar por ciertas razones más bien que por otras, y en que esa decisión es indeterminada. Así, la teoría del agente causal de

²⁸¹ Clarke 1996: 26.

²⁸² Clarke 1996: 24.

Clarke termina siendo una nueva versión del indeterminismo causal. Confío en que no hará falta que reinicie aquí el largo proceso que me llevó, al analizar las teorías de Van Inwagen, Nozick y Kane, a rechazar esta respuesta. Por otra parte volveré a Clarke en breve, para formular contra él una acusación más grave que las precedentes: la de ser un dualista de sustancias.

Pero esto pertenece ya al siguiente apartado, y por ello voy a dar por concluida aquí la presentación de las principales teorías libertaristas que dominan el panorama actual de la filosofía. En las páginas que siguen defenderé que estas teorías en general no hacen sino apostar por patrones libertaristas que a lo largo de la historia de la filosofía se han revelado infructuosos por recurrir a elementos conceptuales que, lejos de aclarar la cuestión de la libertad, la han complicado.

I.1.4. La noción de “causa sui”, el “clinamen” y la tercera antinomia de la razón pura como destinos del libertarismo

Sin duda, cualquier persona que tenga un hermano o una hermana mayor habrá vivido en la infancia una situación parecida a la que voy a relatar brevemente. Nuestra hermana nos llama a su lado y nos dice “mira lo que he aprendido hoy” y entonces, ante nuestra atónita mirada, apaga sin tocarla, tan sólo cubriéndola con un vaso, una cerilla que había logrado permaneciera de pie sobre una mesa. “¡Oh! ¿Cómo lo has hecho?” –preguntamos ingenuamente. Y nuestra hermana contesta con una amplia sonrisa entre pícara, enigmática y satisfecha: “¡Ah! Es magia.” De una forma similar algunos teóricos del libertarismo parecen conformarse, cuando tratan de definir la libertad, con sonreír satisfechos, pícara y enigmáticamente, mientras dicen “¡Ah! Es magia.”

Lamentablemente llega un día en el que crecemos, tenemos a la misma profesora de física y química que nuestra hermana mayor, y nos enseñan lo que es la combustión y cómo ésta es imposible sin oxígeno. Aprendemos entonces que lo que nos habían dicho que era mágico resulta que no lo es, es más, de repente nos damos cuenta de que ya no sabemos muy bien qué quiere decir la palabra “magia”. De niños la entendíamos, o creíamos entenderla, perfectamente: se supone que ciertas personas tenían el poder de obrar milagros. Pero esos supuestos milagros han resultado ser fácilmente explicables, siempre acaba tratándose de algún truco, un truco que, por muy sencillo que sea, a menudo cuesta trabajo descubrir. Con la libertad pasa algo parecido, cuando ya no se confía en cierto tipo de poderes, cuando ya no se es capaz de creer que se entiende qué son algunas cosas que sí se entendían, o se creían entender, en la Edad Media, entonces se pregunta uno muy seriamente si en vez de magia la libertad no será sino un truco. ¿Quiere esto decir que no existe la libertad? En absoluto; quiere decir exactamente lo contrario: la libertad existe tanto que no es ningún milagro; es incluso explicable. Hay quien prefiere seguir diciendo que la cerilla se apaga por arte de magia y tal vez hay alguna razón coherente para hacerlo, pero ésta no puede ser el hecho de que entonces peligre la cerilla. ¿Deja la cerilla de apagarse al explicar científicamente cómo se apaga? No, y parece que la libertad tampoco. Lo que nos enseña nuestra profesora no hace el mundo menos misterioso y apasionante, de hecho siempre es mejor un misterio resoluble que uno imposible de resolver. El mapa de la isla del tesoro tiene gracia si hay tesoro, y así mismo el tesoro sólo tiene interés si podemos encontrar un mapa que nos permita dar con él. Es cierto que lo que sí hace nuestra profesora es sembrar en nosotros un cierto escepticismo acerca de la fiabilidad de futuras explicaciones de nuestra hermana, sobre todo de aquellas en que se haga referencia directa o indirectamente a la magia. En el caso de nuestra hermana tal vez sea una lástima descubrir que es humana... pero en general esa pequeña

sospecha hacia los que todo lo explican por la magia (propia o de un mago más grande y poderoso) puede ser muy beneficiosa.

¿Para qué todo este pequeño preámbulo? Porque no sólo creemos lo que podemos, sino que también tenemos una cierta responsabilidad sobre lo que queremos creer. Creer en algo no es nunca un hecho aislado y sin consecuencias. Más arriba se ha visto en qué fallan las teorías libertaristas actuales más importantes, aquí voy a intentar mostrar que algunas de ellas llevan implícito un cierto compromiso con la magia. Ese compromiso es involuntario en Epicuro, responsable en Agustín y los demás filósofos medievales, sospechoso en Kant, empieza a ser altamente desagradable en Schelling, bien intencionado pero desesperado en Sartre y definitivamente insostenible hoy en día.

¿De qué magia estoy hablando? De la *causa sui*, del *clinamen* y del dualismo de sustancias. Recuérdense una de las citas que han aparecido en mi relatoría de las teorías libertaristas: “(...) el paradigma explicativo de las ciencias naturales en el que las leyes de la naturaleza son las que hacen conexiones explicativas. La fascinación hacia este paradigma puede, parece, cegarle a uno ante el hecho de que el paradigma explicativo de nuestras explicaciones corrientes de la acción mediante razones es bastante diferente.”²⁸³ Es muy fácil llegar a confundir estas palabras de Carl Ginet con estas otras de Schelling: “(...) toda acción singular surge de una necesidad interna del ser libre y que por lo tanto aparece con necesidad, pero una necesidad que al menos no puede ser confundida con una necesidad empírica procedente de una obligación (pero que sin embargo no es también más que contingencia encubierta) (...).”²⁸⁴ Más claro agua: “¡Apagar una cerilla! Eso es coser y cantar, pero apagar un fuego es otra cosa. Lo que dice el profesor sólo sirve para algunas cosas, pero para otras nada puede sustituir a la magia.”

²⁸³ Ginet 1989, en O'Connor 1995: 91.

²⁸⁴ Schelling 1989: 229.

Naturalmente la palabra “magia” está hoy en día muy desprestigiada salvo en la literatura. Muchos magos dirán que no hacen magia, pero tampoco ciencia, sino otra cosa. ¿Pero hay otra cosa? El curandero es un científico disfrazado y el que dice extraer tumores con sus propias manos es un mago disfrazado, un impostor. Schelling, por ejemplo, no es ningún impostor, argumenta inteligentemente para demostrar que no se puede hablar de la libertad *en ciertos términos* si no es presuponiendo un dualismo de sustancias o cuando menos desde fuera del materialismo:

Para poderse determinar a sí mismo, este ser tendría que estar ya determinado en sí, naturalmente no desde fuera, lo que contradiría su naturaleza, ni tampoco desde dentro por alguna necesidad meramente contingente o empírica, en cuanto que todo esto (lo psicológico tanto como lo físico) le subyace, sino que es él mismo, en tanto que su propia esencia, esto es, su propia naturaleza, quien debiera ser su determinación. [...] [S]ólo es libre aquello que actúa según las leyes de su propia esencia y que no se encuentra determinado por ninguna otra cosa ni dentro ni fuera de él.²⁸⁵

Desde mi punto de vista esta idea de Schelling, mal que les pese, se parece mucho (si obviamos el hecho de que ya nadie usa el término “esencia”) a las ideas que manejan muchos libertaristas actuales.

Como hemos visto hay tres posibles formas de hablar de algo así como la libertad con determinación pero sin determinismo (que en el fondo es lo que desean los libertaristas, aunque algunos den a esto el nombre de indeterminismo): bien suponiendo que las acciones surgen espontáneamente del sujeto; bien suponiendo que existen causas que determinan pero que son ellas mismas indeterminadas, y que ellas son el motor de la acción libre; o bien suponiendo que existen dos tipos de causas, unas naturales y otras libres. Si uno es un monista materialista (puede que incluso si uno es un monista idealista) es imposible defender cualesquiera de estas tres opciones. Si uno es materialista todo lo que sucede en el mundo tiene alguna causa anterior y sólo

²⁸⁵ Schelling 1989: 227.

hay un tipo posible de causa, el material. Es cierto que los partidarios del indeterminismo causal suscriben esta tesis, el problema es que no consiguen deshacer el problema del indeterminismo, esto es, el del control del agente. Así, el materialismo no es ni mucho menos por sí mismo una respuesta al problema, sino un punto de partida relativamente inmediato que marca nítidamente los límites de las explicaciones posibles. Ahora, si uno supone el dualismo de sustancias y supone que una de las sustancias es el espíritu y que éste se rige con unas normas diferentes de las de la naturaleza... entonces todo parece funcionar: el espíritu es capaz de causar *ex nihilo*, es más, es capaz de causarse a sí mismo. Si el espíritu es capaz de hacer que haya algo donde no había nada, de crear existencia, entonces es capaz de cualquier cosa y puede explicar el, en palabras de Van Inwagen, “misterio de la libertad”. ¿Pero es esto realmente explicar algo? Bien, aquí no voy a discutir si es coherente la idea de la *causa sui*, sólo querría dejar claro que, aunque de manera encubierta, muchos libertaristas es eso lo que ansían y que Agustín es quien está en el fondo de su teoría. La carga de la explicación está de su lado. En cuanto al dualismo, puesto que existen versiones moderadas y coherentes del mismo, se le puede conceder algo de espacio, siempre que no se trate de un dualismo de sustancias, porque entonces estamos, una vez más, introduciendo una novedad en el mundo que ni es necesaria para entenderlo ni acaba siendo realmente explicativa. Por otra parte, el aceptar el dualismo de sustancias, o incluso uno más moderado de propiedades, y postular que la causación libre es diferente de la causación natural, conduce irremisiblemente a la tercera antinomia kantiana. Es más, la tercera antinomia de la razón pura *consiste* en eso. ¿Alguien quiere una antinomia como punto de partida? Tal vez sea posible llegar a deshacerla, una vez más la carga de la explicación está en manos del libertarista.

Así, como ya avancé en I.1.3, creo que el problema del libertarismo sigue donde lo dejaron Epicuro, Agustín y Kant. El indeterminismo causal puede leerse como una suerte de epicureísmo, una teoría materialista que, para

preservar la libertad contra el determinismo, introduce la posibilidad de que algo no suceda como mandan las leyes de la naturaleza. En cuanto al indeterminismo sencillo, una concepción de la acción según la cual ésta obedece a razones que le son simultáneas y que no tienen antecedente, sino que son producidas espontáneamente por el agente, se parece mucho a la acción de que es capaz el ser humano según la idea que del libre albedrío tiene Agustín, para el que las acciones libres también surgen de la nada. Por último, la distinción en que se basan las teorías del agente causante, entre causas naturales y causas de la acción, no es sino una reformulación de la tercera antinomia kantiana.

Veamos con ejemplos concretos estas semejanzas: en 1. de las semejanzas entre el epicureísmo y el indeterminismo causal, en 2. entre el agustinismo y el indeterminismo simple, y en 3. entre el kantismo y la teoría del agente causante.²⁸⁶

1. Kane señala que su teoría “no postula entidades o relaciones ontológicas adicionales que teorías no libertarias de la libertad de la voluntad tampoco necesitan”²⁸⁷, lo cual es cierto y la hace preferible, desde mi punto de vista, a todas las demás teorías libertaristas. Su teoría es, como la defendida en este trabajo, materialista. Ahora bien, el materialismo por sí sólo no garantiza que no se nos irá la mano con la teoría y acabaremos postulando algo que no aparece fenoménicamente en la realidad. Sin ir más lejos, la teoría de la libertad de los epicúreos era tan materialista como la que aquí se está manejando, y no obstante incluía un incómodo postulado, introducido totalmente *ad hoc* para preservar la libertad: el del *clinamen*. Del mismo modo, Kane nos dice: “La única asunción adicional que he hecho para explicar la libre agencia libertaria

²⁸⁶ He decidido ordenar estos apartados siguiendo el orden histórico, ello conlleva variaciones con respecto al orden seguido en el apartado anterior.

²⁸⁷ Kane 2001b, en Kane 2001a: 237.

no es más que la que cabría esperar —que alguno de los sucesos mentales o procesos involucrados en ellos debe ser *indeterminado*, de tal forma que la causación por sucesos mentales pueda ser no-determinística o probabilística tanto como determinística.”²⁸⁸ Kane incluso admite explícitamente la semejanza de su teoría con la epicúrea: “Como dijeron los antiguos filósofos epicúreos, los átomos deben «desviarse» a caminos indeterminados, si ha de haber sitio para la libertad de la voluntad en la naturaleza.”²⁸⁹ El nombre que recibe el clinamen en la teoría de Kane, en particular, y en las teorías indeterministas, en general, es el de “actividad caótica”. Esta actividad caótica tendría la particularidad de no dejarse notar, como el clinamen, tan sólo la de existir para que quepa la libertad en la naturaleza. En efecto, la indeterminación causal cumple el requisito del agente tan sólo en cuanto entiende que la indeterminación que tiene lugar a nivel microscópico no puede notarse a nivel macroscópico. Mi pregunta es la siguiente: ¿en qué beneficia a las teorías de la libertad algo que no ayuda en absoluto a explicar la acción humana, puesto que es indetectable a nivel macro y es en ese nivel en el que funcionan las explicaciones de nuestras razones para actuar? Parece que tan sólo ayuda a consolar a los preocupados libertaristas, al igual que el clinamen ayudaba a consolarse a Epicuro, aun cuando le planteaba numerosos problemas a su teoría atómica del universo que, por otra parte, era coherente y funcional. Postular el clinamen es menos incómodo que postular una causa sui o un dualismo de sustancias, pero es también más inútil. El caos/clinamen es como el amuleto que lleva alguien que no cree en el poder de los amuletos, pero se ha acostumbrado a llevarlo desde la infancia de tal modo que, aunque no le ayuda llevarlo, su falta le causa de un tremendo desasosiego. En fin, una vez más estoy hablando de magia, en este caso de magia blanca absolutamente inocua, tan inocua que es irrelevante.

²⁸⁸ Ibidem.

²⁸⁹ Ibidem: 238.

2. En lo que respecta al indeterminismo sencillo, la solución de Ginet se parece mucho a la de Agustín. ¿Para qué preocuparnos por el tipo de causalidad que está en el origen de nuestras acciones, si podemos decir que dichas acciones se explican por sí mismas? Si consideramos que cada agente es *causa sui* se disuelve el problema del indeterminismo, ya que por definición cada agente tendrá control sobre sus acciones en tanto será su causa primera y última. Y en efecto así lo ve Agustín: “Por lo cual, ¿qué necesidad hay de investigar de dónde procede este movimiento por el que la voluntad se aparta del bien inmutable para entregarse a los bienes pasajeros, reconociendo, como reconocemos, que es un movimiento propio del alma (...).”²⁹⁰ Pero también lo ve así Ginet, que no encuentra que sea necesario recurrir a la noción de causa (pues ello nos obligaría a buscar antecedentes para la acción del agente) cuando basta, para explicar una acción, con fijarse en la acción en sí misma: “[Las condiciones que explican la acción] no implican nada acerca de lo que, si es que algo lo hizo, causó la acción o cualquiera de los sucesos, físicos o mentales, ingredientes de la acción.”²⁹¹ Además, Ginet y Agustín comparten dosis importantes (casi mortales) de internismo: “(...) la conexión explicatoria [de la acción] es una relación *interna* entre la intención explicante y la acción explicada: se sigue de propiedades *intrínsecas* de los relata. La conexión explicatoria tiene lugar, no por una relación causal, sino simplemente por la referencia directa y la relación interna.”²⁹² No argumentaré más allá de este punto, tan sólo diré que dudo que pueda considerarse a la hipótesis de la *causa sui* y al internismo, buenos compañeros de viaje. No obstante si alguien está dispuesto a admitir ambas ideas no tengo nada que decirle, es su

²⁹⁰ Agustín 1971: 343.

²⁹¹ Ginet 2002: 389.

²⁹² Ginet 2002: 390.

problema, y es él quien debe convencerme a mí de que tales hipótesis son ciertas y que encajan con nuestra experiencia del mundo, y no al revés.

3. “Dándose cuenta de que la libertad de la voluntad no puede ser meramente indeterminismo o azar, [los libertaristas] han recurrido a varias formas oscuras y misteriosas de agencia o de causación para maquillar la diferencia. Immanuel Kant dijo que no podemos explicar la libertad de la voluntad en términos científicos y psicológicos, aun cuando lo requiramos para creer en la moralidad. Para dar razón de ella tenemos que recurrir a la agencia de lo que llamó un «yo nouménico» fuera del espacio y del tiempo que no podría ser estudiado científicamente. Muchos otros filósofos respetables siguen creyendo que sólo cierto tipo de recurso al dualismo mente/cuerpo podría darle sentido a la libertad de la voluntad.”²⁹³ En efecto, es ingenuo pensar que los autores que defienden que las acciones y los efectos son productos *esencialmente* diferentes porque son fruto de dos formas de causación *esencialmente* diferentes, la agencial y la sucesiva respectivamente; es ingenuo, digo, pensar que puede recurrirse a este dualismo en los objetos causados y en el proceso mismo de causación, sin tener que recurrir a su vez al dualismo de los sujetos causantes.²⁹⁴ Se trata de “(...) estrategias del «factor extra». La idea (...) es que, puesto que el indeterminismo deja abierta la cuestión del modo en que un agente elegirá o actuará, algún tipo «extra» de causación o agencia debe ser postulada (...).”²⁹⁵ Y así procede Kant: “(...) la libertad es considerada como un tipo de causalidad que atendiendo a las acciones posibles por su mediación como fenómenos en el mundo sensible, no se halla sometida a fundamentos de determinación empírica (...).” (Ak. V,

²⁹³ Kane 2001b, en Kane 2001a: 227.

²⁹⁴ Tanto O'Connor como Clarke niegan este hecho. Clarke se muestra partidario en principio de un monismo materialista, y O'Connor considera que no le es necesario ir más allá de un dualismo de propiedades. Creo que estos principios no son coherentes con sus teorías, o cuando menos con aquello a lo que aspiran explicar sus teorías.

²⁹⁵ Kane 2001b, en Kane 2001a: 227.

67)²⁹⁶ Pero esto, lejos de permitirle resolver la cuestión, la complica, pues tiene como consecuencia directa la tercera antinomia de la razón pura. Dentro de la filosofía material, opuesta a la formal, Kant señala que puede haber dos tipos de metafísica: la metafísica de la naturaleza y la metafísica de las costumbres. Esta distinción entre disciplinas es resultado de un hiato insalvable entre dos entidades claramente diferenciadas: “El concepto de causalidad como *necesidad natural*, a diferencia del concepto de causalidad como *libertad*, atañe sólo a la existencia de las cosas, en tanto que sea *determinable en el tiempo*, con lo cual no concierne a las cosas como fenómenos, en oposición a su causalidad como cosas en sí mismas. Ahora bien, si se toman las determinaciones de la existencia de las cosas en el tiempo por determinaciones de las cosas en sí mismas (...), no hay manera de conciliar necesidad y libertad en la relación causal, sino que ambas quedan contradictoriamente contrapuestas. Ya que de acuerdo con la primera todo acontecimiento y, por lo tanto, también cualquier acto que transcurre en un punto del tiempo es necesario bajo la condición de que lo era en el tiempo precedente.” (Ak. V, 94)²⁹⁷ La solución que propone Kant a este problema es realmente muy parecida a la de Randolph Clarke²⁹⁸, como puede verse aquí:

Kant: “La necesidad natural, que no puede coexistir con la libertad del sujeto, simplemente se agrega a las determinaciones de aquella cosa que se halla bajo condiciones temporales y, por consiguiente, sólo a las del sujeto que actúa en cuanto fenómeno (...). Más ese mismo e idéntico sujeto (...) considera también su propia existencia como determinable sólo por leyes que se da él mismo a través de la razón, (...) toda acción y en general toda determinación mudable de su existencia acorde con el sentido interno, (...) no son vistas por la consciencia de su existencia inteligible

Clarke: “(...) la relación causal no desempeñaría realmente el mismo papel entre *sustancias*, sucesos y propiedades, que desempeña entre sucesos y propiedades”³⁰⁰

“(...) la ocurrencia de ciertos sucesos previos sería una condición necesaria de la causación de un suceso determinado por el agente. [...] [L]o que causa directamente un agente, cuando actúa con libertad de la voluntad, es que actúa por razones ordenadas de una determinada forma (...)”³⁰¹

“(...) la causa de la acción es el agente, una sustancia, no el suceso que consiste en el causar el agente esa

²⁹⁶ Kant 2000: 153.

²⁹⁷ Kant 2000: p.193.

²⁹⁸ Escojo la versión de Clarke por ser la más sensata, ciertamente no tiene mucho mérito hoy en día enseñarse con Chisholm (con el primer Chisholm). Por otra parte las restantes versiones de la causación agencial (salvo la de O'Connor y otras pocas) son incluso agustinianas antes que kantianas.

sino como una consecuencia de su acción.”³⁰²
causalidad en cuanto *noúmeno* (...).” (Ak. V,
97-98)²⁹⁹

Pero esto no es una solución al problema, es más bien un análisis pormenorizado del mismo, como bien reconoce Kant: “Más todavía nos aguarda una dificultad relativa a la libertad, por cuanto ella debe quedar conciliada con el mecanismo natural en un ser que pertenece al mundo sensible, dificultad que incluso una vez convenido todo lo anterior amenaza con hacer naufragar a la libertad.” (Ak. V, 100)³⁰³ Este problema no es otro que el problema mente/cuerpo expuesto desde unos fundamentos dualistas à la Descartes.³⁰⁴ Así, las teorías que explican la libertad distinguiendo entre el agente como causa y los objetos naturales como causa, acaban teniendo que preguntarse por la glándula pineal. Kant no conseguirá resolver esta cuestión ³⁰⁵, habrá de establecer la libertad como postulado de la razón práctica... pero esta solución no debería dejar tranquilos a los libertaristas, porque precisamente la duda incompatibilista surge en el momento en que uno, al sospechar que el determinismo es cierto, comienza a preguntarse si es correcto postular la libertad. En fin, la solución kantiana sería un petición de principio (¿qué postulado *ad hoc* no lo es?), que como mucho podría satisfacer

³⁰⁰ Clarke 1996: 26.

³⁰¹ Clarke 1993: 204 y 205.

²⁹⁹ Kant 2000: 197-198.

³⁰² Clarke 1996: 24.

³⁰³ Kant 2000: 202.

³⁰⁴ Como ya señalé en la introducción histórica no es obligatorio leer así a Kant (es incluso poco recomendable, a menos que uno sea un idealista a la antigua usanza), no obstante esta lectura es la que más ha pesado entre aquellos autores que se han dedicado al problema de la libertad de la voluntad. Desde luego no es el caso de Rawls, Brandom y otros autores que actualizan a Kant prescindiendo del aparato nouménico.

³⁰⁵ La función de la *Crítica del juicio* sería resolver este problema, y la herramienta conceptual capaz de hacerlo sería el juicio teleológico. Pero el caso es que la clausura del todo razón pura/razón práctica no llega a consumarse en la obra de Kant. Considero que una prueba más o menos inmediata que se puede esgrimir a este respecto, es el hecho de que Kant construyera una *Metafísica de las costumbres*, por un lado, y una *Metafísica de la naturaleza*, por otro. No cabe hacer aquí consideraciones más eruditas acerca de la obra de Kant, porque ello me llevaría demasiado lejos en algo que es accesorio para este estudio. Quien ansíe argumentos más

a un compatibilista, pero desde luego no a un incompatibilista; y en cualquier caso sólo podría satisfacer a un dualista, pues la libertad es posible, según Kant, porque es producto de un noumenon, esto es, porque hay una diferencia entre noumenon y fenómeno. En términos más actuales puede decirse que la libertad es posible porque hay una diferencia entre agentes y causas: en Kant la libertad es la causalidad propia del noumenon, y la necesidad la propia del fenómeno; en O'Connor la libertad es la causación agencial, y la necesidad la causación sucesiva; en Clarke la libertad es la causación por la sustancia que es el agente, y la necesidad la causación entre sucesos. En definitiva, la teoría del agente causante es una propuesta que pretende resolver el problema del indeterminismo postulando un hiato antinómico entre la naturaleza y los agentes. Un hiato que ni Kant, ni los partidarios de la teoría del agente causante, son capaces de salvar.

El punto de vista de este trabajo es el de que lo mejor es lo más sencillo, aplicar el principal principio pragmatista de la historia del pensamiento: la navaja de Ockham. Antes de ponerse a multiplicar los entes hay que asegurarse de que es indispensable. Y parece que no lo es, por ello en los próximos capítulos me voy a dedicar a tratar de construir una teoría coherente y explicativa del *fenómeno* de la libertad. Trataré de fijarme en y de explicar nuestras vivencias de la libertad, no de hacerlas encajar con un *a priori*. Ya no concederé pues, plagiando la expresión de Dennett, más audiencias al libertarismo, tan sólo le cogeré el teléfono de vez en cuando para que me ayude a entender mejor mis propias ideas, pero dando por hecho ya que el libertarismo es una injustificada apuesta por la *causa sui*, el *clinamen* o un descorazonador sendero que empieza y acaba en la tercera antinomia de la razón pura.

sofisticados hará bien en consultar el apartado II.8. “La irradicable tensión del dualismo” en Duque 1998.

Me gustaría creer que detrás de la defensa de algunos tipos de libertarismo no hay un anhelo schellingiano de absolutos, y que si los libertaristas acaban teniendo que postular incómodos dualismos y aún más incómodas generaciones espontáneas no es por nostalgia hacia algún tipo de esencialismo, sino porque realmente ven una importante amenaza para nuestra libertad en el determinismo. Sin embargo, ciertos argumentos que esgrimen los libertaristas son altamente sospechosos. Muchos de ellos, viéndose cercados por algún compatibilista, acaban defendiendo que lo que ocurre es que la verdadera libertad, la de la voluntad espontánea, sólo se da muy de vez en cuando en nuestras vidas. Hemos visto el ejemplo de Van Inwagen, según el cual la libertad sólo tiene lugar en casos de “lucha moral” o de “indecisión existencial”. Así mismo, Kane es un pensador que defiende esta postura con sentido, hablando de acciones puntuales que nos marcan especialmente y nos ayudan a constituirnos como sujeto. Pero hemos visto que otros pensadores como Carl Ginet llegan a plantear situaciones realmente ridículas. Recuérdese que este último pone como ejemplo de verdadera libertad el hecho espontáneo de coger el teléfono con la mano izquierda en vez de con la mano derecha sin saber por qué, sin razón alguna; o también elegir, por supuesto sin razón alguna, escribir un asterisco en el ordenador apretando indistintamente cualquiera de las dos teclas que permiten hacerlo. ¿Qué puede explicar que escojamos una tecla en vez de la otra? Siendo sincero: no lo sé. ¡Ah! Y este “no lo sé” llena, sospechosamente, de gozo al libertarista. Es en ese momento en el que uno se pregunta si realmente lo que le preocupa al libertarista es la libertad, porque parece que es más bien la magia en sí misma, y que tan sólo se ocupa de la voluntad porque en ella le parece descubrir uno de los últimos resquicios del mundo por el que podría colarse la magia. Cuando un filósofo señala como ejemplo paradigmático de la expresión de la libertad la elección de una tecla en vez de otra para escribir un asterisco, algo va mal en esa filosofía. No es tanto que no pueda haber una explicación para ese hecho al

margen de la pura espontaneidad es que... ¿a quién le importa? Tal vez no haya explicación para ese hecho, pero seguramente, aunque la haya, tardaremos en descubrirla porque generalmente a aquellos a quienes les interesa la libertad les preocupa otro tipo de casos que son moralmente más relevantes que la presunta libertad absoluta con la que elijo el asterisco de encima del teclado numérico. Comparto la perplejidad de Daniel Dennett ante las razones que pueden llevar a un filósofo a defender el indeterminismo: “Cualquiera tiene derecho a creer en Gios³⁰⁶ si eso le hace sentir mejor, pero ¿por qué habría de hacerlo? Mi cargo contra los libertaristas es que han convertido unas aspiraciones perfectamente razonables hacia unas versiones valiosas y deseables de la libertad en un anhelo por un tipo de libertad que no sería más deseable ni más valioso que la comunión con Gios.”³⁰⁷

¿Qué ha llevado a los libertaristas a creer en Gios? La respuesta es, sin duda, la contradicción que entre el determinismo y las posibilidades alternativas plantea el argumento de la consecuencia. Así, podría decirse que muchos libertaristas no lo son *motu proprio*, por considerar el indeterminismo como una buena posibilidad en sí misma desde el punto de vista de la libertad de la voluntad (véase Van Inwagen), sino simplemente por rechazo a una tesis, el determinismo, que parece chocar directamente con una de sus intuiciones más básicas con respecto a la acción libre: la de que para que podamos considerar libre a un agente éste ha de ser capaz de elegir entre al menos dos opciones antes de llevar a cabo alguna de ellas.

Así pues, creo que la mejor estrategia para combatir el libertarismo no es tanto la que acabo de desarrollar, esto es, refutar sus argumentos mostrando que se asientan en unas hipótesis extremadamente problemáticas, sino más bien suministrar a los libertaristas una noción de “posibilidad alternativa” que

³⁰⁶ “[U]na gran esfera de cobre que orbita alrededor de una estrella situada fuera de nuestro cono de luz y que lleva prominentemente estampadas en su superficie las letras Gios.”

³⁰⁷ Dennett 2004: 160.

sea coherente con el determinismo, y al mismo tiempo responda a nuestras experiencias más inmediatas de la libertad de acción. Esta tarea es la que me va a ocupar a partir de ahora, con lo que, como ya he dicho, doy por cerrada en este preciso instante la vía libertarista (aunque no podré evitar que, como el río Guadiana, reaparezca de vez en cuando a lo largo del trabajo). Con ello parece que puede considerarse confirmada la hipótesis 1) de la presentación: “Las acciones humanas están determinadas, determinadas por los propios seres humanos.” Puede dar la impresión que ésta tan sólo se ve confirmada por reducción al absurdo, creo que no es así: el determinismo forma parte de nuestra experiencia cotidiana y sólo es puesto en duda por los libertaristas. Por ello, desmontados sus argumentos, en cierto sentido la existencia del determinismo viene dada por el mundo mismo, del mismo modo que, desmontada la duda escéptica, la existencia del mundo no necesita de ningún argumento ulterior para ser aceptada. Queda así salvado el requisito del control del agente, pero se nos presenta el problema del requisito de las posibilidades alternativas.

Hasta este momento he dado por hecho que las posibilidades alternativas no son sólo una condición *sine qua non* de la libertad de la voluntad, sino que son una condición que realmente existe. Ha llegado el momento de poner en duda esa intuición primordial, no tanto porque crea que efectivamente no existen las posibilidades alternativas, sino porque los más hábiles críticos de las mismas son los que han sido capaces de suministrar las más valiosas pistas para concebirlas de una forma cabal y coherente. En el capítulo siguiente voy a ocuparme de esas críticas y de sus réplicas, lo cual me permitirá distinguir entre variedades de las posibilidades alternativas que tienen sentido y variedades que nos empujan a abrazar las abusivas hipótesis metafísicas que en este capítulo me han llevado a rechazar el libertarismo. Voy a intentar descubrir cómo se apaga realmente la cerilla. Al final probaré suerte con mi propia concepción de las posibilidades alternativas. Una concepción

que confío responda a las intuiciones más básicas al respecto pero que, anticipo ya aquí, está muy alejada de lo que se suele llamar “posibilidades alternativas *genuinas*”.

I-2. El jardín de senderos que son ficticios: las posibilidades alternativas dentro de los límites de la mera razón

- (...) El tiempo se bifurca perpetuamente hacia innumerables futuros. En uno de ellos yo soy su enemigo. [...]

- El porvenir ya existe –respondí-, pero yo soy su amigo. ¿Puedo examinar de nuevo la carta?

Albert se levantó. Alto, abrió el cajón del alto escritorio; me dio por un momento la espalda. Yo ya había preparado el revólver. Disparé con sumo cuidado: Albert se desplomó sin una queja, inmediatamente. Yo juro que su muerte fue instantánea: una fulminación.³⁰⁸

Aunque haya que considerar a Robert Nozick el padre de la idea, nadie ha hecho una crítica más potente al principio de las posibilidades alternativas que el filósofo Harry Frankfurt, cuyo artículo “Alternate possibilities and moral responsibility” es un obstáculo al que ningún teórico de la libertad de la voluntad puede dejar de enfrentarse. Se suele considerar, e incluso el propio Frankfurt así lo explica, que su teoría no es una crítica a una determinada forma de entender el requisito de las posibilidades alternativas, sino una enmienda a la totalidad. No obstante, en este trabajo voy a defender que el requisito de las posibilidades alternativas, si bien normalmente se entiende de forma incorrecta, puede tener sentido. Esto no quiere decir que considere que los argumentos de Frankfurt en contra de dicho requisito no son importantes, muy al contrario intentaré elaborar una caracterización de las posibilidades alternativas conciliable con los contraejemplos de Frankfurt. Hay autores que han sido capaces de desarrollar teorías coherentes y funcionales de la

responsabilidad moral al margen de la idea de las posibilidades alternativas (el propio Frankfurt, John Martin Fischer o Daniel Dennett), no obstante considero que esa idea está realmente muy profundamente enraizada en nosotros y forma parte del debate acerca de la libertad de una manera tan esencial, que resulta muy difícil convencer a aquellos que le son favorables de que prescindan de ella, aun cuando los argumentos en su contra sean aplastantes. ¿Quiere esto decir que el requisito de las posibilidades alternativas es incontrovertible por una suerte de esencialismo que lo vincula de forma indiscutible a la libertad? En absoluto, sólo un sentimiento pragmático me anima a defender las posibilidades alternativas: si se puede dar con una definición coherente de dicho requisito, y creo que esto es posible, ello es preferible a simplemente abandonarlo debido a las dificultades que plantea. A dar con esa definición es a lo que dedicaré la segunda parte, *I.2.2.*, de este capítulo intentando confirmar la hipótesis 2) de la presentación. Antes de ello, no obstante, voy a ocuparme en *I.2.1.* de la crítica de Frankfurt al requisito de las posibilidades alternativas.

I.2.1. La crítica a las posibilidades alternativas

En “Alternate possibilities and moral responsibility” Harry Frankfurt sostiene que “(...) el principio de las posibilidades alternativas es falso. Una persona puede muy bien ser moralmente responsable por lo que ha hecho incluso aunque no haya podido hacer otra cosa.”³⁰⁹ Naturalmente, y el propio Frankfurt lo reconoce así, qué deba entenderse por “posibilidades alternativas” está lejos de ser una cuestión exenta de controversia. No obstante Frankfurt cree ver dos corrientes principales en lo que al problema de las posibilidades

³⁰⁸ Borges, “El jardín de senderos que se bifurcan”, en Borges 1971: 115.

³⁰⁹ Frankfurt 1969, en Frankfurt 1988: 1.

alternativas se refiere, fruto de dos formas diferentes de entender lo que es “poder haber obrado de otro modo”: a una de ellas podríamos llamarla una concepción de las posibilidades alternativas en sentido relativo, y a la otra una concepción de las posibilidades alternativas en sentido absoluto.³¹⁰ Frankfurt critica ambas concepciones de las posibilidades alternativas, no obstante cabe pensar, e intentaré mostrarlo en lo que sigue, que las consecuencias de su crítica no son las mismas en uno u otro caso.

En lo que respecta a las posibilidades alternativas en sentido relativo, puede decirse que uno “podría haber obrado de otro modo” si uno no actuó bajo coacción o, por decirlo en términos aristotélicos, en presencia de una fuerza mayor, pero “[l]a idea de que la coacción excluye la responsabilidad moral no es entendida correctamente (...) cuando se toma como una versión concreta del principio de posibilidades alternativas.”³¹¹ El razonamiento de Frankfurt es como sigue: pongamos que un sujeto, Jones, decide hacer algo, pero que alguien le amenaza con un serio castigo si Jones no hace lo que este otro sujeto le pide hacer. Supongamos que el castigo fuera tan aterrador que lo normal sería pensar que cualquier persona razonable obedecería al amenazador, así que Jones le obedece. “¿Consideraremos a Jones moralmente responsable por lo que ha hecho? Pienso que dependerá de los papeles que hayan jugado, a la hora de moverle a actuar, su decisión originaria y la amenaza.”³¹² En efecto, el hecho de que alguien amenace a Jones no implica necesariamente que la acción de éste no pueda ser una acción responsable: ¿Y si Jones actúa de manera compulsiva de tal forma que es irrelevante que se le amenace? ¿Y si Jones se ve forzado a actuar en contra de lo que tenía decidido, por el hecho de la amenaza? ¿Y si el propósito de Jones coincide con el del

³¹⁰ He optado por recoger aquí los términos con los que hablé de diferentes formas de concebir la libertad en la introducción histórica, lo cual espero permita al lector establecer paralelismos entre algunas de las teorías ahí expuestas y la crítica de Frankfurt.

³¹¹ Frankfurt 1969, en Frankfurt 1988: 2.

³¹² *Ibidem*: 3.

amenazador y, aún de no haber mediado amenaza alguna, el resultado habría sido el mismo? Probablemente no sea posible dar una misma respuesta a estas tres preguntas. Así, Frankfurt toma en consideración tres posibles cursos de acción: en todos ellos la voluntad de Jones y la del amenazador coinciden, esto es, lo que el amenazador pretende que haga Jones es exactamente lo que éste tenía pensado hacer de antemano.

Una primera posibilidad es que Jones₁ no sea una persona razonable y que generalmente a la hora de actuar no se deje influir por las consecuencias de sus actos ni por factores externos a su decisión. Así, en este caso la amenaza es irrelevante, si Jones hubiese querido obrar de otro modo lo habría hecho sin importarle lo más mínimo que le amenazasen o no: “Parece evidente que en estas circunstancias el hecho de que Jones₁ estuviese amenazado no reduce de ninguna manera la responsabilidad moral que de otro modo habría soportado por su acción.”³¹³ Lo que ocurre es que Jones₁ es una persona muy especial, así que tampoco sería justo extraer demasiadas conclusiones de su caso.

Un segundo caso posible sería aquel en que Jones₂ se ve absolutamente cohibido por la amenaza. Su pánico es tal que en cuanto es amenazado se le olvida qué es lo que tenía pensado hacer y se somete plenamente a la voluntad del amenazador. Este caso es mucho más interesante que el anterior: “El hecho de que previamente Jones₂ hubiese decidido actuar por sus propias razones precisamente de esa forma podría ser relevante para una evaluación de su carácter; [Jones₂] podría soportar una responsabilidad total por haber tomado *esa* decisión. Pero difícilmente podría decirse de él que fuera moralmente responsable por su acción. Porque llevó a cabo la acción simplemente como resultado de la coacción de la que fue objeto. Su decisión

³¹³ Ibidem.

previa no jugó ningún papel a la hora de causar que hiciera lo que hizo, y sería pues gratuito asignarle un papel en la evaluación moral de su acción.”³¹⁴

Por último, podría darse la situación de que Jones₃ es una persona normal a la que la amenaza podría asustar sustantivamente hasta el punto de hacerle actuar de modo diferente al que él habría preferido, pero dado que lo que se ve obligado a hacer coincide con lo que de antemano tenía previsto hacer, Jones₃ actúa por sus propias razones y no por efecto de la amenaza. “No hay duda de que sería muy difícil para cualquiera saber, en un caso como éste, qué fue exactamente lo que ocurrió. [...] [S]upongamos que nos es evidente que Jones₃ actuó en razón de su propia decisión y no a resultas de la amenaza. Entonces creo que estaría justificado considerar que su responsabilidad moral no se ve afectada por la amenaza aunque, dado que en cualquier caso él se habría sometido a la amenaza, no podría haber evitado hacer lo que hizo.”³¹⁵

El caso de Jones₃ es extremadamente interesante y complicado de analizar en lo que respecta a la atribución de responsabilidad moral. Algunos autores considerarán que Jones₃ estaba realmente coaccionado; otros, entre ellos Frankfurt y el que escribe estas páginas, se decantarían más bien por que no puede hablarse de coacción a pesar de que Jones₃ no habría podido, dada la amenaza, dejar de hacer lo que decidió hacer. Y mi idea es que aquellos autores dispuestos a defender que es un caso claro de coacción lo harían porque le atribuyen una relevancia excesiva al hecho de que Jones₃ sea consciente de la amenaza, cuando en realidad se trata de un aspecto secundario del caso. Se me ocurre el siguiente ejemplo para explicar este punto: pongamos que lo que yo tengo que hacer es elegir entre abrir la caja A o la caja B, y vaciar de su contenido aquella que haya decido abrir. Supongamos que hay alguien, un tal Oswald, interesado en que yo abra la caja B, así que me está vigilando con un

³¹⁴ Ibidem: 3-4.

³¹⁵ Ibidem: 4.

arma de precisión a larga distancia para asesinarme en caso de que yo no decida abrir B. Pero el caso es que Oswald preferiría, si fuera posible, no dejarse ver, y también preferiría, si fuera posible, no tener que asesinarme. Por ello Oswald introduce en la caja A una nota de tal forma que ésta sea lo primero que se vea nada más abrir la caja. Así, si yo optase por abrir la caja A me encontraría con un texto parecido al siguiente: “Tengo una arma apuntándote a la coronilla. Cierra esta caja sin tocar su contenido, abre B y vacíala. Firmado: Oswald.” Supongamos que yo he oído hablar de Oswald, le creo muy capaz de cumplir su amenaza y además tengo razones para creer que no es imposible que Oswald quiera matarme. Pero el caso es que hago que todo el dispositivo de Oswald de eliminación de mi posibilidad alternativa sea inútil porque yo decido abrir la caja B. Creo que Frankfurt estaría de acuerdo conmigo en que es casi imposible hablar de coacción en un caso así. Y el hecho es que la única diferencia que hay entre este caso y el de Jones³ es que éste no ignoraba que estaba amenazado y yo sí, pero este punto parece ser irrelevante porque lo que haría responsable una decisión (ahí parecen apuntar ambos ejemplos) sería el que ésta sea resultado de las razones del agente, que tenga que ver con lo que éste tenía planeado previamente, y no algo que contradiga esas razones a causa de una fuerza mayor. Así, bien podría decirse que uno puede ser responsable moralmente de sus acciones aún bajo coacción, pero yo me inclinaría más bien por decir que en los ejemplos analizados no hay realmente coacción. La clave para explicar este hecho podría estar en la siguiente distinción: uno *no* es responsable de su acción cuando actúa *por* fuerza mayor, y no meramente *en presencia* de una fuerza mayor. Por ejemplo, los gangsters siempre están amenazados, siempre está presente en sus acciones algún tipo de fuerza mayor. Pero seguramente ellos serán capaces de distinguir entre cuando uno de ellos actúa cumpliendo órdenes en una situación normal, o para escapar de un peligro calculado, o dispuesto a sacrificarse y dejarse coger por la policía para evitar males mayores para su jefe; y entre una

situación en que un gangster que lidera una banda rival ha secuestrado a su familia y le amenaza con asesinarla si no se mantiene al margen de un futuro negocio. En fin, fuerza mayor siempre hay en el entorno, para señalar en qué casos ésta es relevante para la responsabilidad hay que evaluar si trastoca o no nuestros planes. Hay una diferencia entre una fuerza mayor que motiva la decisión y otra que no, que es indiferente para la misma: “Supongamos que decidimos que Jones₃ *no* fue coaccionado. Nuestra base para decir esto sería claramente que es incorrecto ver a un hombre como estando coaccionado a la hora de hacer algo a menos que lo haga *a causa de* la fuerza coercitiva ejercida contra él.”³¹⁶ Más adelante (mucho más adelante, en III.3.2.) volveré a ocuparme de esta cuestión con más detenimiento; ahora sigamos con las posibilidades alternativas porque, en cualquier caso, lo que demuestran los ejemplos analizados es que “la doctrina de que la coacción excluye la responsabilidad moral no es una versión particular de las posibilidades alternativas”³¹⁷, pues se puede considerar que los protagonistas de dichos ejemplos no pueden obrar de otro modo pero, o no están realmente coaccionados o, si lo están, aún así no puede decirse que no sean responsables moralmente. “Cuando excusamos a una persona que ha sido coaccionada, no la excusamos porque fuera incapaz de obrar de otro modo.”³¹⁸ Esto no quiere decir que la coacción sea irrelevante a la hora de atribuir responsabilidad moral, tan sólo que no es fiable como criterio para distinguir cuándo actuamos con posibilidades alternativas y cuándo no.

El caso de Jones₃ y el ejemplo que he elaborado en que me veo amenazado por un tal Oswald, bien podrían tomarse como contraejemplos del principio de las posibilidades alternativas. Sin embargo la mayor parte de los autores considerarán que dichos ejemplos sólo contradicen una versión del

³¹⁶ Ibidem: 5.

³¹⁷ Ibidem: 5.

³¹⁸ Ibidem: 5.

principio que no tiene en cuenta las posibilidades alternativas genuinas, pues da la impresión de que dichas posibilidades, para ser relevantes para la libertad, no habrían de ser posibilidades que ofreciera el mundo, sino posibilidades que estuvieran en uno mismo, fruto de alguna “fuerza interna” y no externa. Se trataría de lo que he llamado “posibilidades alternativas en sentido absoluto”. Así, la mayor parte de los defensores de las posibilidades alternativas, especialmente los libertaristas, consideran que la coacción está muy lejos de ser un ejemplo de ausencia de posibilidades alternativas. Para un libertarista, Jones no tiene posibilidades alternativas si es incapaz de querer hacer algo diferente de lo que quiere hacer, por mucho que alguna coacción haga imposible que él haga lo que elija hacer. “En el sentido en que el principio de las posibilidades alternativas usa el concepto de «podría haber obrado de otro modo», la incapacidad de Jones₃ para resistir a la amenaza no quiere decir que no pueda hacer otra cosa que llevar a cabo la acción que lleva a cabo.”³¹⁹ Por ello Frankfurt elabora un cuarto ejemplo, el de Jones₄, que estudié más arriba cuando analicé el indeterminismo sencillo y que conviene recordar aquí:

Supongamos que alguien –digamos que Black- quiere que Jones₄ realice cierta acción. Black está preparado para ir bastante lejos para cumplir sus propósitos, pero prefiere evitar tener que dejarse notar innecesariamente. Así que espera hasta que Jones₄ esté a punto de decidir qué quiere hacer, y no hace nada hasta que tiene claro (Black es un juez excelente en cosas así) que Jones₄ va a decidir hacer algo *diferente* de lo que él quiere que haga. Si se ve claro que Jones₄ va a decidir hacer algo diferente, Black toma las medidas necesarias para asegurarse de que Jones₄ decide hacer, y hace, lo que él quiere que haga. Sean cuales sean las preferencias e inclinaciones previas de Jones₄, pues, Black se saldrá con la suya. [...] [P]ongamos que Black manipula minuciosamente los procesos del cerebro y del sistema nervioso de Jones₄ de una manera directa, de tal modo que las fuerzas causales funcionando dentro y fuera de sus sinapsis y a lo largo de los nervios del pobre hombre determinen que elija obrar y que obre de una forma y no de otra. [En fin,] [d]adas cualesquiera condiciones en que se sostendría que Jones₄ no puede obrar de otro modo (...) pongamos que Black ocasiona el que esas condiciones prevalezcan.³²⁰

¿Por qué este ejemplo es un contraejemplo del principio de las posibilidades alternativas? Porque puede darse el caso de que Jones₄ decida

³¹⁹ Ibidem: 6.

³²⁰ Ibidem: 6 y 7.

hacer algo que está de acuerdo con los designios de Black, y que por lo tanto Black no intervenga en la acción, con lo que Jones₄ estaría obrando de forma absolutamente libre, aún cuando hay que admitir que le sería imposible obrar de otro modo (pues, de darse el caso, Black lo impediría actuando directamente sobre su cerebro). “Sería bastante poco razonable excusar a Jones₄ por su acción, o negarle el halago del que podría hacerle merecedor normalmente, basándonos en el hecho de que no podría haber obrado de otro modo. Este hecho no jugó absolutamente ningún papel a la hora de que él actuase como lo hizo.”³²¹ Al igual que vimos en el caso de la coacción, no es lo mismo el que alguien actúe de una determinada forma sin poder obrar de otro modo, que el que actúe así *porque* no puede obrar de otro modo. Las posibilidades alternativas no son, por sí solas, algo que explique la acción, ni siquiera algo que nos garantice que ésta es realizada en libertad. Nos dice Frankfurt: “El hecho de que una persona no podría haber evitado hacer algo es una condición suficiente para que lo hiciera. Pero, como muestran algunos de mis ejemplos, ese hecho podría no jugar papel alguno en la explicación de por qué lo hizo. Podría no figurar en absoluto entre las circunstancias que realmente provocaron el que hiciera lo que hizo, así que su acción ha de ser juzgada responsable sobre una base enteramente diferente.”³²² Hay circunstancias en que a uno le es imposible obrar de otro modo, pero en que ello es irrelevante porque no es una razón por la que uno decide hacer lo que hace, la acción se explica sin recurrir al hecho de que no había posibilidades alternativas. ¿Por qué concederle entonces a este principio tanta importancia a la hora de atribuir responsabilidad como viene haciéndose tradicionalmente en filosofía desde hace siglos? “A menudo, sin duda, excusamos a la gente por lo que ha hecho cuando nos dice (y la creemos) que no podría haber obrado de otro modo. Pero esto es porque asumimos que lo que nos dice sirve para

³²¹ Ibidem: 7.

³²² Ibidem: 8.

explicar por qué hizo lo que hizo.”³²³ En definitiva, parece que las posibilidades alternativas, *per se*, no garantizan la libertad de una acción, son más bien un adorno deseable. ¿O no?

Seguramente esta conclusión es excesiva. No puede negarse que a menudo recurrimos a la idea de posibilidades alternativas cuando juzgamos nuestras acciones, y tal vez lo que necesita este concepto es más bien una labor de saneamiento que una declaración de edificio en estado de ruina. “Por otra parte, las circunstancias exóticas [como aquellas propias de los contraejemplos de Frankfurt] no ayudan a disipar la ilusión de que, en el curso normal de las cosas, donde no existe tal exceso de determinación, el principio reinante sea, en verdad, el de no considerar responsable a la persona que no se pudo abstener de realizar un acto.”³²⁴ Además, los contraejemplos de Frankfurt han sido sometidos a toda una batería de críticas que han acabado por hacerlos menos inquietantes. A pesar de ello, el partidario de la teoría de Frankfurt, y creo que ello es más que suficiente, siempre podrá seguir defendiendo que “aunque haya alguna posibilidad alternativa de la que el agente es responsable: la posibilidad alternativa no *fundamenta* nuestras teorías sobre la responsabilidad [...]. La cuestión pasa a ser entonces si hay algún *otro factor* que fundamenta nuestras adscripciones de responsabilidad y que también *implica* que hay alguna posibilidad alternativa.”³²⁵ Por ello no me voy a ocupar aquí del interminable debate³²⁶ que han suscitado los ejemplos de Frankfurt, porque en la mayor parte de los artículos que se han ocupado de la cuestión ha quedado claro que cuando una persona nos dice “que hizo lo que hizo porque no podía hacer otra cosa (...). A menudo aceptamos declaraciones como esas (si nos las creemos) como excusas válidas, y declaraciones de ese tipo a primera vista

³²³ Ibidem: 9.

³²⁴ Dennett 2000: 153.

³²⁵ Fischer & Ravizza 1993: 38.

pueden muy bien parecer invocar el revisado principio de posibilidades alternativas. Pero creo que cuando aceptamos declaraciones como esas como excusas válidas es porque asumimos que se nos está diciendo mucho más de lo estricta y literalmente acarreado por la expresión.”³²⁷ Por ello, en el siguiente apartado, me voy a ocupar de intentar desentrañar ese “mucho más” implícito en el “podría haber obrado de otro modo”.

I.2.2. Qué hacemos cuando decimos “podría haber obrado de otro modo”: la forma cotidiana de ser determinista y contar con posibilidades alternativas

Algunos autores hablan de un «sentido incompatibilista de ‘puedo actuar de otro modo’» y un «sentido compatibilista de ‘puedo actuar de otro modo’». Pero cuando los compatibilistas e incompatibilistas castellanohablantes discuten acerca de si la gente podría actuar de otro modo en un mundo determinista, están usando las palabras «podría actuar de otro modo» exactamente en el mismo sentido. De lo contrario no habría desacuerdo alguno, ¿no? [...] Ahora, puede ser que un determinado compatibilista o incompatibilista tenga una *teoría* equivocada sobre aquello que significa «podría haber obrado de otro modo». Pero, en tal caso, ese filósofo no quiere decir ni *él mismo* con «podría haber obrado de otro modo» lo que su errónea teoría dice que significan esas palabras. [...] Entonces ese [filósofo] no está tan sólo equivocado acerca de lo que los demás quieren decir con «podría haber obrado de otro modo»; también está equivocado acerca de lo que *él* dice con esas palabras.³²⁸

Aquí voy a ser tan ambicioso como para pretender que, efectivamente, “podría haber obrado de otro modo” sólo tiene un sentido, pero que éste no es en el que inmediatamente piensa la gente, por usar la expresión de Schopenhauer, “filosóficamente virgen”, ni tampoco es aquel en el que inmediatamente piensa la gente ‘filosóficamente pervertida’. Me atrevería a defender que la interpretación que aquí voy a presentar, aunque en algunas ocasiones ha sido tímidamente esbozada, no ha sido decididamente defendida. La razón por la que los incompatibilistas no habrían sido capaces de hacerlo,

³²⁶ Para hacerse una idea suficiente de las líneas principales de este debate consúltese la parte IV “Moral Responsibility, Alternative Possibilities, and Frankfurt-Style Examples” de Kane 2002: 281-334.

³²⁷ Frankfurt 1969, en Frankfurt 1988: 9-10.

³²⁸ Van Inwagen 1989, en O’Connor (ed) 1995: 222-223.

atendiendo a lo que vengo estudiando, sería obvia: el incompatibilismo es el resultado mismo de una mala comprensión del sentido en que se dice habitualmente “podría haber actuado de otro modo”. La razón por la que los compatibilistas no habrían dado con ella se me escapa pues está presente, aunque de manera implícita, en algunas teorías (concretamente en algunas páginas de Dennett y de Schopenhauer, que no en balde serán mis guías en este capítulo). Tal vez la mayor parte de los compatibilistas no han sido capaces de convencerse a sí mismos de su coherencia, o tal vez la han repudiado por un sano rechazo, bastante generalizado hoy en día, a fiarse de una experiencia fenomenológica de carácter introspectivo, que es el tipo de experimento que me ha permitido dar con ella. Comparto esa suspicacia, no obstante creo que una metodología de este tipo no conduce irremisiblemente al internismo y por ello aquí voy a recurrir brevemente a ella. Así pues voy a analizar primero cuál es la interpretación canónica del principio de las posibilidades alternativas y voy a intentar mostrar su falsedad y sobre todo su esterilidad como principio moral. Pasaré entonces a explicar mi propia teoría e intentaré, en la medida de lo posible, mostrar su coherencia y su aplicabilidad a casos reales de atribución de responsabilidad. Por último intentaré ganar adeptos para mi definición de las posibilidades alternativas comparándola con la tradicional, y mostrando los problemas que presenta desde el punto de vista normativo optar por una teoría de la libertad que pretenda defender un sentido absoluto de aquello en que consiste contar con posibilidades alternativas. Adentrémonos ya, pues, en el resbaladizo territorio de averiguar qué contiene la caja “posibilidades alternativas” que cada uno de nosotros, como la caja del escarabajo de Wittgenstein, lleva consigo.

¿Qué quiere decir un libertarista cuando dice que alguien actuó libremente porque “pudo haber obrado de otro modo”? Generalmente quiere decir lo siguiente: una persona, si es libre, es capaz de, en *circunstancias específicas*, optar entre elegir hacer al menos dos cosas diferentes. Las cursivas señalan

cuál es la parte polémica de esta lectura, que es la más común, de lo que es contar con posibilidades alternativas. Así, contar con posibilidades alternativas no es, según la tradición, ser capaz de hacer una cosa en una ocasión y, para otra ocasión posterior similar a la primera, enmendarse y obrar de modo diferente. No, contar con posibilidades alternativas *genuinas* es tener la posibilidad de, en un mismo momento *t*, hacer A o hacer B, donde A y B son dos acciones que no pueden ser llevadas a cabo simultáneamente. Curiosamente esta versión libertarista de las posibilidades alternativas coincide, según Schopenhauer, con las intuiciones más inmediatas que uno pueda tener acerca de aquello en lo que consiste ser capaz de actuar de otro modo: “(...) el hombre ingenuo (esto es, el filosóficamente virgen) opina que en un caso determinado le fueron posibles actos de voluntad contrarios, apoyándose para ello en el testimonio de su conciencia de sí.”³²⁹ La sentencia de Schopenhauer para con aquel que posee esta intuición es condenatoria: “Confunde desear con querer. Puede desear cosas contrarias, pero querer sólo una de ellas; cual sea ésta se lo revela a la autoconciencia la *acción*.”³³⁰ Una vez más el punto de vista defendido en este trabajo coincide con el de Schopenhauer porque ¿qué podría querer decir que uno tuvo la posibilidad en *un momento concreto* tanto de hacer A como de hacer B? He aquí una posible explicación, la más antigua, la más libertarista, la del *liberum arbitrium indifferentiae*: “Una completa indiferencia en Dios es una prueba muy grande de su omnipotencia.”³³¹ En efecto, la libertad “genuina” es la de la indiferencia, pues desde ésta *todo* son posibilidades absolutamente indeterminadas, son un acto puro de voluntad que nada hace necesario, ningún pensamiento ni ningún motivo, son pura elección. Se trata de una voluntad que no tiene que lidiar con el mundo. En ese sentido *al mismo tiempo* es posible tanto A como B, porque no hay nada anterior

³²⁹ Schopenhauer 2000: 61.

³³⁰ Ibidem.

³³¹ Descartes 1996: 280.

que pueda condicionar una de esas dos posibilidades, ninguna razón que dar a favor de cualquiera de ellas. Pero parece que la libertad a la que podemos aspirar los seres humanos no es de ese estilo: “(...) la indiferencia no pertenece en absoluto a la esencia de la libertad humana, visto que no somos solamente libres cuando la ignorancia de lo bueno y lo verdadero nos vuelve indiferentes, sino principalmente también cuando el conocimiento claro y distinto de una cosa nos empuja a y nos compromete en su búsqueda.”³³² En definitiva: lo propio de la libertad humana es el actuar por alguna razón, esto es, desde la determinación. Pongamos que soy una persona que adora a su novia, para mí ella es lo más importante de mi vida y lo último que querría en este mundo es verla sufrir. Bien, imaginemos que se ve arrastrada por un torrente en un río, yo no soy un nadador experto y la corriente es fuerte y peligrosa, hay pocas posibilidades de que la salve si salto al agua, sin embargo, como no hay nada para mí más importante en el mundo que ella, me lanzo al agua y tengo tan buena fortuna que la salvo. La pregunta es, si a lo que se aspira es a una libertad en sentido absoluto, ¿puede decirse en este caso de manera absoluta que yo tuve la posibilidad de no arrojarme al río, que no arrojarme al río era una verdadera opción en sentido absoluto? Creo que no. ¿Quiere eso decir que no tiene sentido imaginar que yo no hubiese saltado a rescatar a mi novia, que no tiene sentido contemplarlo como una posibilidad real? No, siempre que no se pretenda que se trataba *desde un punto de vista absoluto* de una posibilidad real... lo cual la convertiría meramente en una posibilidad lógica. Ahora voy a explicar esta aparente paradoja.

“(...) [N]uestras acciones son el producto de dos factores, de los que uno, nuestro carácter, permanece firmemente inalterable y sólo nos es conocido paulatinamente <a posteriori>; pero el otro factor lo constituyen los motivos, que moran en el exterior y se ven ocasionados necesariamente por el

³³² Ibidem.

curso del mundo, los cuales determinan el carácter dado, bajo la presunción de su estructura estable, con una necesidad asemejable a la mecánica.”³³³ Este cuadro es el que suele aterrar a los libertaristas, y probablemente debería inquietarnos a todos a menos que seamos capaces de renunciar a la libertad absoluta que Descartes le atribuye a Dios y que los libertaristas reclaman para sí. También, el argumento de la consecuencia nos hace concebir pocas esperanzas para nuestra libertad si el cuadro schopenhaueriano es acertado. Bien, desde mi punto de vista ese cuadro es tosco pero correcto. Efectivamente mi “carácter” y mis “motivos”, esto es, mis genes pero también mi educación y todas las demás influencias culturales externas me condujeron a tener, en el momento en que mi novia cayó al agua, un determinado “estado cerebral” en sentido lato, traducible a un motivo racional, que me impulsó a lanzarme al agua y rescatar a mi novia. “(...) [N]uestros padres y maestros no ignoran que sin una educación moral nos convertiremos en adultos egoístas, indignos de confianza y probablemente peligrosos. De manera que se las ingenian para educarnos y, gracias a sus esfuerzos responsables, somos capaces de apartar de nuestra mente todo pensamiento de latrocinio, engaño o violencia. Estas alternativas nos resultan impensables en circunstancias morales; además nos enseñaron a cuidarnos a nosotros mismos y a contribuir a nuestro desarrollo moral -¿acaso ser una persona responsable no consiste, en buena medida, en inhabilitarse para cometer acciones que *nos volverían culpables?*-.”³³⁴ Dados mis genes y, de modo especial, mi educación, no me fue posible obrar de otro modo que como obré salvando a mi novia porque en ese momento no podía querer otra cosa que la que quería. ¿No obré entonces libremente? Sí, porque si volviese a darse un caso *casi* idéntico (por definición, al ser un nuevo caso, no podría ser absolutamente idéntico) sería posible (en este caso posible realmente) que yo actuase de modo diferente, puesto que ya

³³³ Schopenhauer 1994 : 20.

³³⁴ Dennett 2000: 155.

habría recibido nuevas influencias externas que podrían haber alterado mi querer. Ahora, ¿es que cuando decimos “obrar de otro modo en *exactamente las mismas circunstancias*” el tener el motivo que acaba por impulsarnos y el que éste se imponga sobre las alternativas, no forma parte de esas “mismas circunstancias”? “Exactamente lo mismo” significa que no hay nada diferente. ¿Cómo puede nadie pretender seriamente que dos causas exactamente idénticas en exactamente las mismas circunstancias produzcan efectos diferentes? Para que ello fuese posible sería necesario el continuo milagro que postula Malebranche, y ni aún así podría hablarse de “circunstancias exactamente iguales” pues no sería lo mismo el milagro que me lleva a hacer A que el que me lleva a hacer B.

En el momento en que es evidente, que el hombre no es libre de querer o no, la primera cosa que se pregunta después de esto es, si el hombre es libre de querer cuál de los dos prefiere; ¿el movimiento por ejemplo o el reposo? Pero esta cuestión es tan visiblemente absurda en sí misma, que puede bastar para convencer a quienquiera que reflexione sobre ella, de que la libertad no concierne en ningún caso a la voluntad. Ya que preguntar si un hombre es libre de querer, qué prefiere, ¿el movimiento o el reposo, hablar o callarse? es preguntar, si un hombre puede querer lo que quiere, o preferir lo que prefiere, cuestión, que desde mi punto de vista, no necesita respuesta.³³⁵

Así pues, si contar con posibilidades alternativas significa tener poder para hacer A o B en un mismo y único momento, entonces no contamos con posibilidades alternativas. No obstante, creo que sí contamos con ellas a pesar de que cada vez que actuamos no podemos hacer otra cosa que la que en efecto hacemos.³³⁶ “«Casual» significa la coincidencia en el tiempo de algo que no es unido causalmente. Pero nada es *absolutamente* fortuito; sino que incluso lo más azaroso no deja de ser algo necesario, sólo que llegado por un camino

³³⁵ Leibniz, *Nouveaux Essais sur l'Entendement Humain* pf. 25, en Leibniz 1959: 255.

³³⁶ En este punto de la investigación parece que podría resultar útil trabajar ya con una definición de posibilidad que evite malentendidos, creo no obstante que aún es posible continuar manejando una idea intuitiva de posibilidad. Me ocuparé detalladamente de este problema en *I.3.1.*, pero me gustaría sugerir aquí como telón de fondo conceptual para los párrafos que vienen el apartado 194 de las *Investigaciones filosóficas* de Wittgenstein, donde habla del error que supone entender las posibilidades de la máquina como algo que preexiste en la máquina (sobre este tema considero muy recomendable la lectura de *Significado y libertad* de César Gómez). Lamento no poder

muy lejano; pues en algún punto remoto de la cadena causal, se hallan las causas que ya han determinado necesariamente, hace largo tiempo, que tal cosa tuviera lugar justo ahora y a la vez que aquella otra.”³³⁷ Naturalmente esto implica, si el argumento de la consecuencia es correcto, que no podemos ser libres, porque las leyes de la naturaleza y los sucesos del pasado remoto no dependen de nosotros. Pero aquí sigo empeñado en afirmar que sí somos libres, sólo que no en sentido absoluto. Y parece que va llegando el momento de explicar más exactamente qué quiero decir con eso de “en sentido absoluto”. Recurriré una vez más a las certeras palabras de Schopenhauer: “(...) a nadie se le ocurre (...) disculparse de una mala acción apelando a esa necesidad (...). Pues se da cuenta de que esta necesidad tiene una condición *subjetiva* y que, *objetivamente*, esto es, en presencia de las mismas circunstancias, actuando los mismos motivos que le determinaron a él, hubiera sido posible una acción distinta y hasta contraria, si únicamente *él hubiera sido otro*.”³³⁸ Dicho de otro modo: aquel que aspira a la libertad en un sentido absoluto está dando por sentado que en un mismo momento una misma persona es al mismo tiempo dos personas diferentes, puesto que puede querer dos acciones contrarias al mismo tiempo. Quien aspira a la libertad que efectivamente existe se da cuenta de que, siendo uno quien es, sólo puede actuar como de hecho actúa, pero que *se puede hablar de acción libre cuando ésta tiene lugar en circunstancias en las que cualquier otra persona podría haber obrado de otro modo*, y no se puede hablar de acción libre cuando las circunstancias son tales que, aún cuando en sentido absoluto existían otras posibilidades, el no haber actuado como se actuó no parece razonable o moralmente exigible. Sencillamente una cosa es lo lógicamente posible y otra lo empíricamente posible. Creo que hay una diferencia entre elegir sin ningún tipo de presión si se quiere pintar las paredes

extenderme más al respecto, confío en que seré capaz de llenar más adelante el hueco que dejo aquí.

³³⁷ Schopenhauer 1994: 31-34.

de casa en verde o en azul, o elegir con una pistola en la sien. Sin embargo, si nos empeñamos en hablar de la libertad en sentido absoluto ambas situaciones se hacen idénticas: “(...) si decimos que alguien, queriendo mentir para no perder la vida o la salud, se ve obligado a traicionar la verdad por temor de la muerte o de los tormentos, no decimos la verdad, porque no está más obligado a querer la vida que la verdad; pero como le está prohibido, por una fuerza extraña, conservar una y otra a la vez, *escoge* él mismo lo que ama más, *espontáneamente y no contra su voluntad* (...)”³³⁹ Si realmente esta visión de Anselmo es la versión más intuitiva que hay de la libertad, resulta difícil de entender la existencia de nuestros actuales sistemas legislativo y judicial. Tal vez se trate de la versión correcta, pero por algún extraño ardid intelectual no del todo bienintencionado se nos ha llegado a convencer de que también es la más inmediata, y eso contradice frontalmente nuestra experiencia diaria de la libertad y la responsabilidad. Más bien parece que sea la versión más cristiana, lo cual, en una sociedad como la nuestra tal vez se haya convertido, es cierto, en algo bastante inmediato... pero no por ello más intuitivo. La magia no es más intuitiva que el conocimiento, sólo es más inmediata. En cualquier caso, desde un punto de vista absoluto de la libertad los seres humanos, *independientemente de las circunstancias*, siempre somos plenamente libres de elegir lo que queremos, y es precisamente esa irrelevancia de las circunstancias lo que hace que, en unas *circunstancias específicas* como conllevarían las posibilidades alternativas genuinas, podamos elegir hacer A o B. Más adelante (en II.4.) desarrollaré un argumento pragmatista para rechazar este tipo de consideraciones ya que, como puede verse, la libertad entendida en sentido absoluto no nos reporta ningún bien de carácter normativo, no nos permite hacer toda una serie de distinciones que son indispensables para el juicio moral. De momento voy a usar un argumento fenomenológico para intentar

³³⁸ Schopenhauer 2000: 152.

³³⁹ Anselmo de Aosta 1952: 579.

dotarnos de una teoría del significado de “podría haber obrado de otro modo” alternativa a la canónica. Sumerjémonos pues en nuestra propia experiencia de las posibilidades alternativas, en el acaecer de la oración “podría haber obrado de otro modo” en nuestra vidas. ¿Qué pensamos realmente cuando decimos “podría haber obrado de otro modo”?

En primer lugar hay que distinguir entre dos posibles sujetos de ese pensamiento. “Podría haber obrado de otro modo” puede ser pensado en primera y en tercera persona. En un caso se trata de alguien que, *de forma retrospectiva* analiza su acción y, dado que no le gustan sus consecuencias, se imagina tomando una decisión diferente de la que efectivamente tomó y actuando así de modo diferente a como efectivamente obró. En el otro caso se trata de alguien que analiza, también de manera *retrospectiva*, la acción de una tercera persona y se ve a sí mismo en las mismas circunstancias en que aquella persona actuó y eligiendo llevar a término una acción diferente de la que dicha persona realmente llevó a cabo.³⁴⁰

Para empezar es importante notar que en ambos casos el pensamiento se da de forma retrospectiva. Este hecho está lejos de ser inocente. Mi hipótesis es que el tipo de juicio implícito en el “podría haber actuado de otro modo” necesita, para poder ser construido, tener en cuenta las consecuencias de la elección del agente del que sea cuestión. Así, yo pienso de mí mismo que podría haber obrado de otro modo, y probablemente lo deseo, a la vista de las circunstancias posteriores a mi acción, pues es un hecho (como se ha mostrado más arriba) que no tiene sentido pensar que yo efectivamente podía

³⁴⁰ Se puede encontrar alguna excepción a este patrón. Por ejemplo, un juez no se “ve a sí mismo” actuando en lugar del acusado. Este hecho se explica, y seguramente suceda algo parecido con otras excepciones que quepa citar, porque en el caso de un juicio está especificado el método que ha de seguirse para atribuir responsabilidad. La ley viene a sustituir el tener que implicarse en la experiencia del agente (aunque también un juicio puede leerse como un proceso de reconstrucción objetiva y pormenorizada de las circunstancias del agente y, en este sentido, aunque nuestra mirada sobre los sucesos deba ser distante, se nos hace participar de la experiencia del acusado).

elegir no querer hacer lo que en el momento de mi acción quería hacer. Por otra parte, en tercera persona, yo puedo verme a mí mismo habiendo evitado lo que era inevitable, siendo el otro y actuando de forma diferente a como él actuó.

Así, “podría haber obrado de otro modo” en primera persona viene a ser un “de haberlo *sabido* habría hecho A en vez de B”, y en tercera persona viene a ser “de haber estado en *esa* situación *yo* habría hecho A en vez de B”. En ambos casos se hacen hipótesis salvajes (imposibles de corroborar), pero nos pasamos la vida haciendo hipótesis de este tipo. ¿Basta con esto para hablar del requisito de las posibilidades alternativas? Pienso que sí, y para justificar esta afirmación voy a intentar mostrar que ambos modelos son coherentes viendo si son capaces de dar cuenta de algunas excepciones al comportamiento responsable, esto es, voy a observar si las recién formuladas redefiniciones del “podría haber obrado de otro modo” eximen de responsabilidad a aquellas personas a las que normalmente, en el día a día, estaríamos dispuestos a eximir de responsabilidad.

Consideremos el caso de una persona que, en un momento de enajenación mental, asesina a otra. Tenderíamos a considerar que no es plenamente responsable de su acción (aunque este punto deberá ser tratado con más cuidado más adelante, concretamente en III.3.). ¿Es coherente la imagen de esa persona pensando “podría haber obrado de otro modo”? No, y por ello tendemos a eximirle de responsabilidad. Y ahora la prueba: ¿es coherente la imagen de esa persona pensando “de haberlo sabido no le habría matado”? Tampoco. Es exactamente igual de incoherente ya que la enajenación mental supone un eximente porque es un estado que hace irrelevante la deliberación, o que incluso constituye una ausencia de deliberación, esto es, una situación en la que no puede caber aquello que “de haberse sabido” podría habernos hecho cambiar nuestro curso de acción. Así pues parece que esta versión de las posibilidades alternativas en primera

persona es coherente con nuestras intuiciones acerca de aquello en lo que consiste el poder de actuar de otro modo, veamos ahora la definición en tercera persona.

Realmente a menudo juzgamos poniéndonos en la piel del otro, o poniéndonos en su lugar, pero nunca imaginándonos que *somos* el otro, pues resulta obvio que si nos considerásemos exactamente iguales que el otro, dadas exactamente las mismas circunstancias, actuaríamos igual que él.³⁴¹ Pero no, la gracia del juicio moral es que nos comparamos y pensamos que, si para nosotros habría sido posible no haber cometido determinado acto, entonces dicho acto no era inevitable en un sentido moralmente relevante, por mucho que dicho acto fuese, desde el punto de vista de las leyes naturales y la causalidad universal, inevitable para la persona que lo llevó a cabo, dados su “carácter” y sus “motivos”. Consideremos un nuevo caso de exención: un hombre secuestra a la pareja y a los hijos de una mujer y, bajo la amenaza de asesinarlos si ella no hace lo que él le pida, la obliga a robar a sus propios padres. Supongamos que a ella le es imposible acudir a la policía o reclamar la ayuda de cualquier otra persona, así que finalmente roba a sus padres. ¿Encontramos responsable a la mujer? Desde luego que no porque consideramos que, aunque desde un punto de vista absoluto ella podría no haber actuado como actuó, no obstante ceder a la coacción de ese infame secuestrador era razonable y que probablemente “de haber estado en su lugar habríamos hecho lo mismo” o, dicho de otro modo, consideramos que, dadas las circunstancias, no había posibilidades alternativas relevantes.

Seguramente esta caracterización de las posibilidades alternativas no satisface a muchos libertaristas, así como a todos aquellos que son

³⁴¹ Un experimento mental que corrobora este punto puede verse en Corbí 2003. Josep Corbí habla de unos virtujuegos en que “el experienciador podría elegir sus rasgos psicológicos con el fin de descubrir cómo vive el mundo una persona diferente”, lo cual, en el fondo, es una ilusión porque “¿qué quedaría entonces de él, pues lo que vale para el carácter también se aplica a los deseos, proyectos, compromisos, ilusiones, etc del experienciador en cuestión?” (pp. 19 y 20).

agustinianos pero no lo saben; no obstante parece ser más relevante desde el punto de vista moral la versión que de las posibilidades alternativas se ha construido aquí, que la que acaba haciéndole admitir a Anselmo que cuando alguien miente para conservar la vida es responsable de ser un mentiroso. Parece que la versión aquí defendida tiene fuerza normativa, es coherente y además responde al acaecer fenoménico del pensamiento “podría haber obrado de otro modo” en nuestra vida cotidiana. “[C]uando enjuicia un acto para culpar, aunque la gente efectivamente pregunta «¿Podría haber hecho otra cosa?», no muestra interés en llevar esa cuestión más allá del punto en que han satisfecho su curiosidad acerca de la existencia o ausencia de fatalismo local [circunstancias que no permiten escoger libremente tales como la coacción o la compulsión] —un fenómeno que es enteramente neutral respecto al determinismo y el indeterminismo.”³⁴² Así pues, considero que puede decirse que los seres humanos, salvo excepciones que ya tendremos ocasión de estudiar, son seres libres que cuentan con posibilidades alternativas a pesar de estar determinados, o más bien gracias a que están determinados. Ese contar con posibilidades alternativas supone que *un agente, cuando actúa, podría haber actuado de otro modo, esto es, de haber sabido cuáles iban a ser las consecuencias de su acción tal vez no la habría llevado a cabo, o bien cualquiera, o al menos la mayor parte de la gente, en su lugar habría o podría haber actuado de otro modo*. Creo que con esto es suficiente para la libertad y la responsabilidad moral, aunque necesita de un desarrollo casuístico que tendré que dejar para la segunda y tercera partes de este trabajo.

Dos problemas subsisten no obstante: nuestras intuiciones más inmediatas sobre la libertad (o tal vez no tan inmediatas, sino fruto de siglos de agustinismo), y el argumento de la consecuencia. De momento voy a intentar disipar las primeras, de lo segundo me ocuparé en el siguiente capítulo.

³⁴² Dennett 1984, en Kane 2001a: 85.

Por una parte mucha gente puede pensar que cuando dice de alguien que “podría haber obrado de otro modo” se está refiriendo a las posibilidades alternativas genuinas, a que esa misma persona en esas mismas circunstancias podría haber obrado de otro modo. Es más, da la impresión de que dicho pensamiento no siempre se da de forma retrospectiva, sino que a menudo, y en eso consiste el proceso de deliberación, pensamos “puedo hacer A o puedo hacer B”. Seguramente esto es cierto y somos capaces de pensar algo así como que *si hubiésemos querido*³⁴³ otra cosa *entonces* la habríamos hecho, pero esto no es más que señalar una obviedad y resulta normativamente estéril.³⁴⁴ Lo que no resulta estéril es ponernos en lugar del otro con nuestros propios motivos, comparándolos con los de ese otro y dándonos cuenta de que tener unos motivos y no otros era relevante para una acción dada porque no mediaban ni la coacción, ni la compulsión, ni cualquier otro factor que normalmente consideramos suspende nuestra responsabilidad. Sin embargo, como a pesar de estas consideraciones seguirá habiendo quien defienda que cuando dice “podría haber obrado de otro modo” no pretende en absoluto decir “de haberlo sabido no lo habría hecho” o “yo en su lugar no lo habría hecho”, voy a intentar mostrar de forma persuasiva que decir cualquier otra cosa no tiene sentido. Como se ha visto, la caracterización tradicional de las posibilidades alternativas pretende que dos cosas opuestas sean posibles al mismo tiempo. A

³⁴³ También es interesante considerar la posibilidad de que de haber tenido *creencias* diferentes, entonces seguramente también habríamos podido obrar de otro modo. A este respecto me remito a Thiebaut 2000b. En dicho artículo Carlos Thiebaut lleva a cabo un análisis fenomenológico de un acto comunicativo similar al “podría haber obrado de otro modo”, a saber, el de “ponte en mi lugar”. Thiebaut explica que llevar a cabo esa operación, la de ponerse en el lugar del otro, supone adoptar su horizonte de razones, constituido por sus creencias, intenciones y supuestos normativos. Aquí hablo únicamente de motivos, como aquello que mueve directamente a la acción y que presupone ese horizonte razones.

³⁴⁴ Esta solución compatibilista al problema de las posibilidades alternativas fue propuesta originariamente por Moore (aunque ya aparece en Hobbes y en Hume) y después adoptada por numerosos defensores del compatibilismo. En general ha sido abandonada porque conlleva un problema de regresión al infinito: a la respuesta “podría haber hecho otra cosa de haberlo querido”, se puede contestar con otra pregunta “¿podría haber querido otra cosa?”, a lo que cabe responder “de haberlo querido habría querido otra cosa”, que a su vez implica una nueva pregunta “¿podría haber querido querer otra cosa?” y así hasta el infinito (“y más allá”, Buzz Lightyear dixit).

este respecto Daniel Dennett nos insta a considerar el caso de Austin fallando un golpe fácil jugando al golf y pensando después “(...) me enfado conmigo mismo porque podría haberlo embocado. (...) estoy hablando de las condiciones exactamente como se dieron, y afirmo que podría haberlo embocado.”³⁴⁵ Dennett propone que hagamos pasar esta afirmación por una prueba en términos de mundos posibles: “Primero, supongamos que el determinismo es verdadero y que Austin falla; sea H la proposición «Austin emboca el golpe». Debemos escoger ahora el conjunto X de mundos posibles relevantes que debemos sondear para ver si podría haberlo conseguido. Supongamos que escogemos como X el conjunto de mundos físicamente posibles que son *idénticos* al mundo presente en algún tiempo t_0 previo al golpe. Como el determinismo afirma que en cada momento hay exactamente un solo futuro físicamente posible, dicho conjunto de mundos tiene un solo miembro, el mundo presente, el mundo donde Austin falla.”³⁴⁶ “Claro, -responderá el libertarista- pero es que si uno escoge como conjunto de mundos posibles el conjunto de mundos físicamente posibles idénticos a este...” Pero, ¿acaso no habla Austin en este ejemplo de “las condiciones exactamente como se dieron”? ¿Y es que un mundo diferente de este no es una condición muy diferente? Los libertaristas deberían darse cuenta de que, cuando hablan de “exactamente las mismas circunstancias” ni siquiera el indeterminismo los puede salvar. Supongamos que arrojar una moneda al aire (lo cual no es realmente cierto pero nos servirá para este ejemplo) sea un sistema indeterminista de elección. Pongamos que decido que si sale cruz voy al cine y si sale cara doy un paseo por el Parque del Retiro. Pongamos que sale cruz y voy al cine. ¿Qué sentido tiene decir que, en exactamente las mismas circunstancias, podría haber sucedido que hubiese salido cara y hubiese ido al Retiro? Ese “exactamente las mismas circunstancias” supone también la

³⁴⁵ Citado en Dennett 2004: 96.

³⁴⁶ Dennett 2004: 96-97.

ocurrencia del lanzamiento de la moneda y de su resultado. Los “exactamente” los carga el diablo.

No tiene sentido que dos cosas sean posibles en t .³⁴⁷ Las posibilidades alternativas como mucho consisten en que antes de la acción (en este sentido puede decirse “puedo hacer A o puedo hacer B”), en t_1 , es posible hacer A porque en t_1 al agente se le ocurre o quiere hacer A pero no llega a hacerlo; y en t_2 es posible hacer B porque al agente se le ocurre o quiere hacer B pero no llega a hacerlo; y es posible que en t_3 el agente se esté debatiendo entre A y B. En este sentido, obviamente tanto A como B fueron posibles, aunque no *exactamente* al mismo tiempo. Cuando llega el tiempo t que corresponde a la decisión definitiva y a la acción, entonces ya sólo A o sólo B es posible y real al mismo tiempo. La posibilidad simultánea se daría tal vez en el tiempo t_3 posterior a la ocurrencia de las posibilidades aunque anterior a la decisión definitiva, esto es, en el momento de la deliberación. Tal vez pueda decirse que en ese momento en cierto sentido se contaba con posibilidades alternativas simultáneas aunque, una vez más, no en sentido absoluto. Tal vez muchos preferirían quedarse con esta versión de las posibilidades alternativas en vez de con aquella que he construido, sin embargo este tipo de formulación de las posibilidades alternativas dota de sentido, lo cual desde mi punto de vista la hace claramente insatisfactoria como teoría, a lo siguiente: si a Ramírez en t_3 , cuando podría tener la posibilidad de matar o de no matar a Carreño, no se le ocurre otra posibilidad que no sea matar a Carreño, entonces en t , cuando efectivamente mata a Carreño, Ramírez no es libre de no matar a Carreño, pues no existía ninguna posibilidad alternativa en t_3 que fuera no matar a Carreño. Se puede bautizar a este ejemplo como “el problema del asesino poco ocurrente”. En efecto, en este caso nuestro Ramírez es tan tonto que

³⁴⁷ Me remito aquí a la nota 336 y véase también un poco más adelante. Insisto en que en *I.3.I.* me ocuparé de definir “posibilidad” como es debido. Insisto también en que, si posible, se lea este apartado con un ojo wittgensteiniano.

sólo se le ocurre una posibilidad, y si hacemos depender la libertad del hecho de que haya varias, entonces no somos capaces de considerar responsable a Ramírez de matar a Carreño. Si subordinamos la libertad a la existencia de posibilidades en el pensamiento del agente, el juzgar a alguien moralmente responsable depende demasiado de la imaginación del agente y de su facilidad para crear mentalmente estados de cosas posibles.³⁴⁸ Por ello parece que vincular la libertad a la existencia real de posibilidades abiertas en el mundo es no entender correctamente la libertad: “(...) diferentes conjeturas sobre lo que hay no definen *diferentes posibilidades*.”³⁴⁹ Hay un supuesto internista muy peligroso en la idea de las posibilidades alternativas genuinas en tanto se refiere a los motivos del agente como mundos posibles que éste imagina antes de actuar. “Pensar de esta forma es lo que nos lleva a la extraña idea de que en la regla preexisten, de alguna manera, sus aplicaciones, que la interpretación de la regla está ya en la regla, como los movimientos posibles están ya, en algún sentido, en la máquina. Esta manera de pensar no es metafísicamente aséptica, ni mucho menos, sino que se enraíza con toda una tradición modal que ha sustancializado las posibilidades, las viejas potencias aristotélicas.”³⁵⁰ Si las posibilidades alternativas deben darse *realmente* en el agente, deben suceder *en él* (aquí está el presupuesto internista), entonces la libertad depende en exceso de la imaginación de cada uno.

De todas formas, como se ha podido ver con el tiro al hoyo de Austin, nadie tiene realmente posibilidades alternativas en exactamente un mismo instante t . Defenderlo es imposible a menos que, como acaba de ser sugerido, convirtamos t en un momento alargado formado por t_1 , t_2 y t_3 , y no en un instante. Pero entonces nos basta con quedarnos con la intuición de que

³⁴⁸ Una importante característica de la responsabilidad es su carácter *recursivo*: es posible atribuir responsabilidad a alguien independientemente de que éste fuese consciente o no de que estaba haciendo un mal. Para aprender más sobre este aspecto de la responsabilidad véase Thiebaut 1998.

³⁴⁹ Gómez 2003: 85.

³⁵⁰ Ibidem: 107.

somos libres si podemos deliberar, independientemente de que sólo se nos ocurra un curso de acción posible. Y aún esto tal vez sea excesivo, pues hay ocasiones en que no tenemos tiempo de deliberar y simplemente actuamos: a los soldados se les entrena para ello, ¿quiere esto decir que no son libres? En absoluto, no es necesario que deliberemos, sino que seamos *seres deliberantes*, con esa capacidad. No todos los seres humanos (como se verá más adelante en III-4.) han sido correctamente contruidos como seres deliberativos y por ello pueden ser eximidos de responsabilidad. No obstante la mayor parte de los seres humanos son capaces de deliberar, piensan moralmente y se dotan de principios que se graban fuertemente a sí mismos para asegurarse de que en ciertas situaciones serán capaces de prescindir de la deliberación y actuar rápidamente pero de forma coherente con sus deseos y creencias. Una vez más eso no quiere decir que al actuar inmediatamente estén actuando sin deliberación, lo que pasa es que ésta tuvo lugar tiempo atrás y su resultado estaba ya previsto, el agente decidió automatizar sus respuestas ante ciertos desafíos que sabía no le iban a permitir entretenerse en hacer ciertas consideraciones (¿no es así como se convierte uno en un ser virtuoso según Aristóteles?). Por todo ello, creo basta con la forma de caracterizar normativamente las posibilidades alternativas aquí defendida, que sí da cuenta de esas situaciones y que puede resumirse, tanto para la primera como para la tercera persona, del siguiente modo: “otro, en mi/su lugar, podría haber obrado de otro modo”, donde ese “otro” puede entenderse como uno mismo pero con diferentes motivos (seguramente a la luz de unas consecuencias inesperadas de la acción).³⁵¹ Ciertamente no nos expresamos así, pero no creo que sea errar en la interpretación de la experiencia de la atribución de responsabilidad, pensar que lo que hacemos es colocarnos en la misma situación que un determinado agente e imaginar qué haríamos o no haríamos.

³⁵¹ Diferentes motivos que pueden ser resultado de diferentes creencias, deseos... En fin, cualquier cosa relevante para la acción que sea susceptible de cambiar a la luz de sus consecuencias.

No nos imaginamos siendo quienes no somos, sólo nos imaginamos viviendo la situación que otro vivió, en su lugar físico, ir más allá es imposible y si sí lo hacemos seriamente, entonces nos damos cuenta de que lo propio sería actuar igual que aquel agente a quien estamos sustituyendo.³⁵² Somos capaces de ponernos en la piel de otro, pero seguimos siendo nosotros, es imposible dejar de ser uno mismo y ser otro, incluso mediante la imaginación. Lo único en que consisten las posibilidades alternativas es en tener ocasión de deliberar y que esa deliberación *sea relevante para la elección de mi acción*, o cuando menos que la acción se corresponda con principios fruto de deliberaciones anteriores que me permitirían haber incorporado ciertos patrones de acción-respuesta, pero “[c]onsiderar ambos lados de una cuestión no significa que uno no sea finalmente persuadido por uno de ellos.”³⁵³ Cuando alguien me apunta con una pistola y me obliga a hacer algo, puedo deliberar, ahora, mi deliberación sólo afecta de forma débil a mi acción, ésta no será genuinamente mía. Desde el punto de vista absoluto es innegable que tengo libertad de acción, pero ese no es el sentido de la libertad relevante desde el punto de vista moral porque, como ya ha sido señalado y como se verá con más detalle en las partes II y III de este trabajo, no es el que nos permite hablar de responsabilidad moral.

Nosotros sugerimos que Austin haría mejor en considerar que «Austin emboca el golpe» es posible si, en situaciones *muy parecidas* a la situación en cuestión, Austin emboca el golpe. Pensamos que esto es lo que quería decir en realidad, y que haría bien en conceptuar de este modo su golpe. Ésta es la manera familiar, razonable y útil de conducir los «experimentos ulteriores» siempre que estemos interesados en comprender las causas que intervienen en un fenómeno que nos interesa. Variamos levemente (y a menudo metódicamente) las condiciones

³⁵² La idea de Carlos Thiebaut es que este “convertirse” en el otro no sólo no es posible, sino que ni tan siquiera responde a la demanda de aquel que nos pide que nos pongamos en su lugar, pues éste pretende que juzguemos su situación (por ello Thiebaut habla de un uso “judicatario” del “ponerse en el lugar del otro”) haciendo *epojé* de nuestro horizonte de razones pero sin con ello dejar de ser nosotros. Así, la petición “ponte en mi lugar” incluye al menos tres demandas: “[1] Normalmente, cuando se nos pide «ponte en mi lugar» no se nos solicita que compartamos la totalidad de las posibles creencias de quien nos habla (...) sino un conjunto de creencias *relevantes* para el caso. [...] [2] «Ponte en mi lugar» parece también incluir, por tanto, algo así como una cláusula «dado como yo entiendo esta circunstancia y supuestos mis propósitos que forman parte de ese mi entendimiento... ponte en mi lugar». [...] [3] «Ponte en mi lugar» indica, por lo tanto, que juzguemos la acción desde un sistema de *pertinencias* de normas y principios, seleccionando entre todos los posibles aquéllos que el interlocutor juzga adecuados al caso.” Thiebaut 2000b, en Bertomeu, Gaeta & Vidiella 2000: 33, 34 y 35.

³⁵³ Dennett 2000: 154.

iniciales para ver qué cambia y qué permanece inalterado. Ésta es la forma de reunir información *útil* del mundo para guiar nuestras campañas ulteriores de evitación y progreso.³⁵⁴

En lo que respecta al argumento de la consecuencia alguien podría decir: “Bien, hasta aquí se me ha convencido de que sólo una forma de entender las posibilidades alternativas tiene sentido, pero ello no me demuestra mi libertad sino todo lo contrario, pues no da respuesta al argumento de la consecuencia; sólo demuestra que la primera premisa del argumento, esto es, el determinismo, es correcta y que por lo tanto yo no soy libre, y que precisamente por eso no cuento con posibilidades alternativas genuinas.” No se puede, pues, soslayar por más tiempo una respuesta al argumento de la consecuencia, y a ello voy a dedicar el primer apartado del próximo capítulo. Llegó pues el momento de concluir este.

La conclusión es que no se puede tomar la libertad en un sentido absoluto porque entonces pierde su significado el concepto de posibilidades alternativas que usamos con sentido y coherencia en nuestra vida cotidiana, pero la libertad es ante todo un concepto normativo, que cobra sentido por su uso. Tal vez sea difícil persuadir a la gente de que a menudo usa “podría haber obrado de otro modo” de forma errónea, pero parece una empresa titánica (de la que trataré en II.3.) convencerla de que no es libre y responsable, así que prefiero elegir la primera tarea, más modesta. Por otra parte, si las consecuencias de defender las posibilidades alternativas “genuinas” son, como muestra crudamente la cita de Anselmo de Aosta recogida más arriba, que uno siempre es libre y responsable sin importar las circunstancias, entonces uno puede empezar a dudar seriamente de que las posibilidades alternativas de las que solemos hablar cotidianamente sean “genuinas”. Normalmente no atribuimos responsabilidad preguntándonos por si era lógicamente concebible un querer diferente en el agente del querer que se llevó el gato al agua. Realmente nadie diría, como sí lo hace Anselmo, que puesto a elegir entre su

³⁵⁴ Dennett 2004: 98.

vida y la mentira uno es libre de escoger no mentir. Y tampoco solemos pretender averiguar, cuando preguntamos a un especialista en un juicio sobre si el acusado “pudo haber obrado de otro modo”, si las circunstancias físicas que redundaron en la acción que debemos juzgar estaban o no determinadas: “Si (...) nuestra responsabilidad depende de la pregunta «¿podríamos haber obrado de otra manera *exactamente* en aquellas circunstancias?», nos enfrentamos con un problema de ignorancia de lo más peculiar (...). [E]s extremadamente improbable, dada la complejidad del cerebro aun en el nivel molecular (...) que podamos probar de manera contundente que un acto particular es el efecto en gran escala de una indeterminación en el nivel subatómico crítico. En consecuencia, (...) es muy probable que nunca tengamos ninguna razón para creer que un acto particular fue o no responsable. [Y] quienes dicen saber que podrían haber obrado de otro modo exactamente en aquellas circunstancias deben admitir que proclaman este hecho, presumiblemente empírico, sin el menor atisbo de evidencia y sin la más remota esperanza de obtener tal evidencia.”³⁵⁵ La libertad interesa como concepto para resolver situaciones de conflicto, no en sí misma como entelequia. Si optamos por una definición de libertad como libertad absoluta, la discusión está fuera del mundo y carece de interés, “[s]i queremos comprender lo que ocurre en ese mundo [de agentes determinados], podemos –o en realidad debemos– hablar sobre los cambios que introducen [los agentes] en sus circunstancias a través de sus elecciones informadas, y sobre lo que *pueden* y *no pueden* hacer. [...] Mirar *exactamente* el mismo caso, una y otra vez, no resulta nada informativo: lo que lleva en realidad al diagnóstico es prestar atención a otros casos parecidos.”³⁵⁶ De ese modo aprendemos los seres humanos, viendo algo que no nos gusta y pensando “yo no haré lo mismo cuando esté en su situación”, lo cual presupone, acertadamente, que *diferentes*

³⁵⁵ Dennett 2000: 156-157.

³⁵⁶ Dennett 2004: 102-103.

personas, dado que tenemos capacidad deliberativa, somos capaces de actuar de manera diferente en situaciones parecidas. Y esto último no es sino la hipótesis 2) de la presentación, a saber, “una o más personas pueden actuar de forma diferente en situaciones muy parecidas”, que creo haber sido capaz de corroborar en párrafos anteriores. Dicha hipótesis es la que corresponde a mi particular lectura del principio de las posibilidades alternativas, y confío en que haya quedado demostrada en este capítulo su vigencia en nuestra vida cotidiana, frente a la vigencia de aquella otra versión más que humana de las posibilidades alternativas que pretende ser la más inmediata e intuitiva pero que está muy lejos de serlo: “(...) el interés por saber si se podría haber obrado de otra manera, exactamente en las mismas circunstancias y en el mismo estado interno, supone la existencia de una curiosidad metafísica particularmente pura, tan pura que carece de todo motivo ulterior, ya que la respuesta no modificará en nada el transcurrir del mundo.”³⁵⁷ No deberíamos olvidar que la libertad del libertarista, incluso del libertarista actual como se pudo descubrir en apartados anteriores, es una libertad absoluta, en último término divina, que pretende darse en este mundo físico pero atendiendo a las normas del mundo de las posibilidades lógicas y no de lo empíricamente posible. Pero la libertad que debería interesarnos es la humana, pues sólo en el contexto de nuestras prácticas morales cotidianas acaece la libertad y cobra sentido la pregunta por las posibilidades alternativas. Y sobre esa libertad tan humana, demasiado humana dirán algunos, se va a trabajar aquí:

Ya que [nuestra libertad] consiste solamente en que podemos hacer una cosa o no hacerla (...); o más bien solamente en que, para afirmar y negar, ir detrás o huir de las cosas que el entendimiento nos propone, actuamos de tal forma que no sentimos en absoluto que ninguna fuerza exterior nos constriña. Pues, para que sea libre, no es necesario que sea indiferente a elegir entre uno u otro de los dos contrarios; sino más bien, tanto más que me incline hacia uno (...). De manera que esa indiferencia que siento, cuando no soy llevado más bien hacia un lado que hacia el otro por el peso de ninguna razón, es el grado más bajo de la libertad, y hace más bien aparecer un defecto en el conocimiento que una perfección en la voluntad; pues si conociese siempre claramente lo que es verdadero y lo que es bueno, nunca

³⁵⁷ Dennett 2000: 158.

me vería forzado a deliberar acerca de qué juicio y qué elección debería hacer; y así sería enteramente libre, sin ser nunca indiferente.³⁵⁸

I-3. Edipo no podía ser sino Edipo... pero el oráculo no lo sabía: la libre voluntad y el agente

Ahora me encuentro que soy un miserable y nacido de miserables. ¡Ah, encrucijada de doble senda y selvoso valle oculto, encinar y desfiladero en un cruce de dos caminos, que la sangre de mi padre que es también la mía por obra de mis manos bebisteis, ¿os acordáis aún de mí, qué clase de acciones llevé a cabo ante vosotros y luego, al venir aquí, cuáles hice de nuevo?³⁵⁹

—¡Oh Ahab! —gritó Starbuck—. No es demasiado tarde para desistir, ni aún en el tercer día. ¡Ya ve, *Moby Dick* no le busca! ¡Es usted quien la busca a ella insensatamente!³⁶⁰

En *I.1.4.* me pregunté qué es lo que lleva a los incompatibilistas a abrazar hipótesis como la de la causa sui, el clinamen o el dualismo de sustancias. Ya ahí anticipé una posible respuesta: el argumento de la consecuencia, según el cual si es cierto el determinismo, entonces no hay posibilidades alternativas. En el capítulo anterior he tratado de mostrar que es posible hablar de posibilidades alternativas desde el compatibilismo, ahora voy tratar de deshacer el argumento de la consecuencia. Con ello espero disminuir la necesidad de los libertaristas de recurrir a rebuscados giros metafísicos para dar cuenta de la libertad. Tras esto aventuraré muy esquemáticamente una explicación genealógica de por qué los libertaristas acaban, como dije en *I.1.4.* plagiando a Dennett, por creer en Gios. También trataré de confirmar la hipótesis 3) de la presentación, esto es, que los seres humanos predecimos comportamientos pero ello no implica que seamos seres perfectamente predecibles.

³⁵⁸ Descartes, René, 1996: 88-89.

³⁵⁹ Sófocles, *Edipo Rey* (1390-1400), en Sófocles 1988: 287.

³⁶⁰ Melville, *Moby Dick II*, en Melville 2004: 365.

Esta parte se cerrará con un análisis del sujeto responsable, esto es, con un estudio de la conciencia que aúna las acciones que, además de ser libres, son coherentes con una serie de razones y que por ello responden a un carácter o personalidad que las hace, por citar la distinción de Rawls, racionales y razonables, y por ende *responsables*. Este análisis es pues necesario para que en la segunda parte del trabajo nos adentremos en el problema de la responsabilidad moral. Por desgracia sólo en los últimos capítulos del trabajo será posible completar el cuadro del agente responsable que comenzaré a esbozar aquí. Muchos elementos que están más allá del problema metafísico de la libertad de la voluntad han de ser estudiados antes de que podamos entender el carácter dual del agente responsable. Aquí me ocuparé de una mitad de su personalidad: de lo que se ha llamado su “verdadero yo” o “carácter”. Pero antes enfrentémonos una vez más al argumento de la consecuencia.

I.3.1. Necesidad y contingencia: determinación no es predeterminación

¿Qué es exactamente el determinismo? Creo que esta definición de Daniel Dennett es muy acertada: “(...) *en cada momento hay un único futuro físicamente posible*. Decir que el determinismo es verdadero es decir que nuestro mundo forma parte de un subconjunto de mundos que poseen la interesante propiedad siguiente: no hay dos mundos que comiencen exactamente igual (si comienzan siendo iguales, siguen siendo iguales para siempre: no son mundos *diferentes* en ningún sentido), y siempre que dos mundos compartan *cualquiera* de sus descripciones de estado, compartirán también todas las descripciones de estado posteriores a ésta.”³⁶¹ Parece pues que, efectivamente, en principio el determinismo es incompatible con la libertad, cuando menos con las

posibilidades alternativas. Como ya se ha visto esta última afirmación no tiene por qué ser cierta, pues cabe concebir las posibilidades alternativas de una forma compatibilista. Pero muchos libertaristas no son capaces de conformarse con este “no tiene por qué ser cierta”, necesitan que se les asegure que es falsa. Hacerlo viene a ser demostrar que el argumento de la consecuencia es incorrecto, o cuando menos no infalible.

Recordemos cómo es el argumento de Van Inwagen:

(1)	$\Box (Po \ \& \ L) \supset P$	
(2)	$\Box (Po \supset (L \supset P))$	Lógica estándar
(3)	$N(Po \supset (L \supset P))$	Regla (α)
(4)	NPo	Premisa
(5)	$N(L \supset P)$	Regla (β)
(6)	NL	Premisa
(7)	NP	Regla (β)

Lo cual, como vimos, puede traducirse del modo siguiente:

- (1) Nuestras leyes naturales y nuestro pasado implican que P necesariamente.
- (2) Si es nuestro pasado el que es, entonces, si son nuestras leyes naturales las que son, entonces P necesariamente.
- (3) Si es nuestro pasado el que es, entonces, si son nuestras leyes naturales las que son, entonces P sin que ello dependa de nosotros.
- (4) El pasado no depende de nosotros.
- (5) No depende de nosotros el que las leyes naturales impliquen que P.
- (6) Las leyes naturales no dependen de nosotros.
- (7) P no depende de nosotros.

Hay bastantes conceptos en el argumento de la consecuencia que necesitan ser aclarados. Así, en (1) “implican que P necesariamente” conlleva

³⁶¹ Dennett 2004: 88.

una cierta definición de la causación (en tanto que P se da como resultado de sucesos anteriores y de la acción de las leyes naturales sobre los mismos) que está lejos de poder ser aceptada sin más como coherente con las consecuencias que Van Inwagen extrae de su argumento. También, en (3) “P sin que ello dependa de nosotros” conlleva una definición de “posibilidad” (no nos es posible hacer que $\neg P$) que no actúa en el mismo plano que la noción de causación presupuesta en (1), con lo que ambas no podrían medirse en el mismo argumento.

En efecto, puede hacerse una lectura contrafáctica de la causación que no equivale a la implicación necesaria que utiliza Van Inwagen, y dicha lectura sería coherente con el determinismo. Supongamos que X es un conjunto de mundos similar al nuestro, que ϕ es el antecedente de una conexión causal y ψ su consecuente. Bien, la siguiente proposición (A) es válida si y sólo si en cualquier mundo más o menos similar al nuestro, cuando se da el antecedente, también lo hace el consecuente:

$$(A) \quad \phi \Box \rightarrow_x \psi$$

Si X cambia, esto es, si tomamos otro conjunto de mundos posibles no cercanos a X, entonces ya no funciona el operador $\Box \rightarrow$ en (A). Pero Van Inwagen recurre al operador \supset o \Rightarrow , que vale para todo mundo posible.

Por otra parte, el determinismo concierne a las condiciones suficientes para que algo se dé, no nos dice nada acerca de sus condiciones necesarias: “si σ_0 es una proposición (increíblemente compleja) que especifica con todo detalle el estado del universo en t_0 y similarmente σ_1 especifica el universo en un momento posterior t_1 , entonces el determinismo dicta que σ_0 es suficiente para σ_1 en todos los mundos físicamente posibles. Pero el determinismo no

nos dice nada acerca de qué condiciones previas son *necesarias* para producir σ_1 , o en realidad cualquier otra proposición ψ . Así pues, en tanto que la causación generalmente presupone necesidad, la verdad del determinismo tendría poco que ver con la validez de nuestros juicios causales.”³⁶² En efecto, la existencia del paramecio (σ_0) en algún momento temprano de la evolución es causa suficiente de que yo ahora esté escribiendo mi tesis doctoral (ψ), pero yo no diría que el que yo ahora esté escribiendo “dependa de” o haya sido causado por la existencia del paramecio. Y eso a pesar de que, según la definición de determinismo vista más arriba, dada la existencia del paramecio en las condiciones exactas en que se dio en este mundo en el momento t_0 , y dado que yo, en el momento t_1 , y en este mismo mundo, estoy escribiendo mi tesis, no podría ser que yo no estuviese escribiendo mi tesis ahora. “Los incompatibilistas que aseguran que bajo el determinismo σ_0 «causa» o «explica» ψ no dan con el objetivo de la investigación causal.”³⁶³ Uno de esos incompatibilistas es Van Inwagen, cuyo argumento puede ser reconstruido del modo siguiente:

1. Sea ϕ un suceso que efectivamente tiene lugar en la vida del agente A (...). Sea también σ_0 una amplia descripción del estado del universo en un momento del pasado remoto, y sea λ una afirmación de las leyes de la naturaleza.
2. Entonces, asumiendo el determinismo, $\lambda \wedge \sigma_0 \Rightarrow \phi$ se aplica a todo mundo posible. También, $\neg\phi \Rightarrow \neg(\lambda \wedge \sigma_0)$.
3. Si A tiene poder para causar α y $\alpha \Rightarrow \beta$ es verdadero en todo mundo posible, entonces A tiene poder para causar β .
4. Así pues si A tiene poder para causar $\neg\phi$, entonces A tiene poder para causar la falsedad tanto de λ como de σ_0 , lo que es absurdo.
5. Por lo tanto A carece de poder para causar $\neg\phi$.³⁶⁴

La clave del argumento está en ‘3.’, donde se afirma que el “poder para causar” es poder para causar necesariamente, y no contrafácticamente. Si en

³⁶² Taylor & Dennett 2002, en Kane 2002: 271.

³⁶³ Ibidem: 272.

³⁶⁴ Ibidem: 273-274.

lugar de \Rightarrow , Van Inwagen usase $\Box \rightarrow_x$, su argumento no se sostendría. Pero es que precisamente la causalidad que implica el determinismo es contrafáctica y no necesaria, con lo que el postular el determinismo tal y como lo hace Van Inwagen en (1), esto es, como $\Box (Po \& L) \supset P$, no es correcto. “Nuestras leyes naturales y nuestro pasado implican que P necesariamente” no describe el determinismo, sino la necesidad lógica que atañe a la verdad de los enunciados. Pero el determinismo es una teoría acerca de la conexión causal de los sucesos del mundo, no sobre la verdad de las aserciones. Una descripción correcta del determinismo sería “nuestras leyes naturales y nuestro pasado son causa suficiente de que P”, y este postulado no conlleva ninguna contradicción con el que ciertas cosas dependan de nosotros contrafácticamente. Si entendemos “poder para causar” como que “A tiene *poder para causar* α si y sólo si para una proposición γ que describe una acción de A y un mundo f cercano a la realidad, $\gamma \wedge \alpha$ es verdadero en f y $\alpha \Rightarrow \gamma$ en todo mundo similar a f .”³⁶⁵ Esto viene a ser lo mismo que decir $\alpha \Box \rightarrow_f \gamma$, esto es, que el poder para causar de A es contrafáctico, y por lo tanto compatible con el determinismo. “En otras palabras, dentro de un conjunto de racimos cercanos, hay una acción posible de A (llamada γ) que es una condición necesaria para que ocurra α . Pero según esta definición, ‘3.’ no tiene garante alguno. ‘3.’ plantea como hipótesis que $\alpha \Rightarrow \beta$ en cualquier mundo; si pudiéramos deducir $\beta \Rightarrow \gamma$ en ese racimo, habríamos alcanzado la meta. Pero (...) $\alpha \Rightarrow \gamma$ y $\alpha \Rightarrow \beta$ no implica $\beta \Rightarrow \gamma$, con lo que ‘3.’ falla, y con él el argumento de Van Inwagen.”³⁶⁶

Bien, una vez elaborado este contra-argumento, dado que presumiblemente el argumento de la consecuencia responde a nuestras intuiciones más inmediatas sobre el determinismo, será útil para la discusión

³⁶⁵ Ibidem: 274.

³⁶⁶ Ibidem.

adoptar un punto de vista más intuitivo, pues es en ese punto de vista donde está en juego realmente la posibilidad de disipar las dudas que sobre el compatibilismo arroja dicho argumento. Proseguiré pues con los problemas que conlleva el concepto de posibilidad dejando ya de lado el argumento de Van Inwagen: “(...) la verdad o la falsedad del determinismo no debe afectar a nuestra creencia de que ciertos eventos no realizados eran sin embargo «posibles», en un *sentido importante y cotidiano* de la palabra.”³⁶⁷ Ese sentido, es cierto, no se corresponde con el de que sea posible, *al mismo tiempo*, hacer y no hacer algo. Sobre el sinsentido que supone hablar de posibilidades alternativas en esos términos ya he escarbado en capítulos anteriores así que no insistiré más en ello. Lo que me interesa ahora es señalar que “[t]anto en nuestras reflexiones cotidianas sobre lo que vamos a hacer a continuación como en nuestro pensamiento científico más riguroso en relación con las causas de los fenómenos, empleamos conceptos de *necesidad*, *posibilidad* y *causalidad* que son estrictamente neutrales respecto a la cuestión de si la verdad está del lado del determinismo o del indeterminismo.”³⁶⁸ Este punto es importante porque existe una enconada discusión acerca de si el indeterminismo de la mecánica cuántica es algo que tiene lugar en los procesos causales de la naturaleza, o si tan sólo se refiere a nuestra facultad cognoscitiva para determinar las evoluciones de esos procesos. Me inclino más bien por la hipótesis de que la indeterminación se refiere a nuestra imposibilidad de conocer de forma no probabilística ciertos sucesos, pero, dada la incesante efervescencia de argumentos en uno y otro sentido, voy a evitar esta cuestión referida de forma concreta a la teoría cuántica y me voy a centrar en análisis más abstractos acerca de algunos de los diferentes significados que le podemos atribuir a las palabras “posible”, “necesario” y “contingente”. En cualquier caso, me parece importante destacar el hecho de que cuando hablamos normalmente de

³⁶⁷ Dennett 2004: 98.

³⁶⁸ Ibidem: 84.

posibilidades no estamos suscribiendo una teoría concreta sobre el funcionamiento del universo, y que si lo hacemos, entonces ésta es determinista. Por mucho que avance la mecánica cuántica, seguirá sin ser útil para explicar la práctica totalidad de los acontecimientos que se dan a nivel macroscópico en la corteza terrestre. En ese ámbito recurrimos a leyes causales de carácter determinista: con ellas explicamos la acción, en ellas basamos la responsabilidad. Dicho esto, una vez aceptado que vivimos en un mundo en que todo sucede según procesos de causa-efecto, me dedicaré a tratar de comprender qué es lo posible, qué lo necesario y qué lo contingente.

Cabe distinguir, básicamente, entre dos formas de entender la posibilidad: como posibilidad lógica y como posibilidad física. Mi idea es que el libertarista, para admitir que existe la libertad, quiere las dos al mismo tiempo, pero esto no es posible (ni lógica, ni físicamente). Por “posible” en sentido lógico, se puede entender “aquello que no es autocontradictorio”. Por ejemplo: un animal terrestre del tamaño de una ballena es posible en sentido lógico porque el concepto de “tamaño” y el de “animal terrestre” no son contradictorios. Así, lo que es verdadero lógicamente, lo es en todo mundo posible, y viceversa. Ahora, “posible” en sentido físico, puede leerse como “aquello que, de hecho, dadas unas leyes naturales concretas y unas ciertas propiedades de la materia, existe o podría existir de darse las condiciones necesarias para ello”. Así: un animal terrestre del tamaño de una ballena no es posible porque los músculos de un animal de ese volumen nunca podrían ser capaces de soportar el peso correspondiente fuera del agua.³⁶⁹ Se ve pues que, a diferencia de lo posible lógicamente, lo físicamente posible está condicionado por *un* mundo, por lo real. Pero precisamente algo que es posible lógicamente es algo que vale para *cualquier* mundo. Los libertaristas

³⁶⁹ Por cada talla que aumente un animal su peso crecerá tres veces, pero la resistencia de sus músculos y huesos sólo dos. La ballena es capaz de soportar su peso gracias al empuje del

mezclan ambos conceptos de posibilidad: por ello hablan de cursos de acción alternativos posibles simultáneamente, porque la posibilidad lógica no entiende de espacio ni de tiempo; y por ello pretenden que esos cursos de acción posibles realmente estén en nosotros a la hora de elegir, porque la posibilidad física no entiende de existencias que no sean materiales (aunque bien es cierto, como señala acertadamente Fernando Broncano en *Saber en condiciones*, que “la materia ya no es lo que era”). Pero esta cordial entente de la lógica y la física es una quimera en lo que al tema de las posibilidades se refiere. Si las posibilidades alternativas son lógicamente posibles no estamos señalando más que una tautología, si las posibilidades alternativas son físicamente posibles no podemos defender que existan simultáneamente. Es posible lógicamente lo que no es contradictorio en cualquier mundo, es posible físicamente lo que no contradice las leyes físicas de este mundo, y ambos conjuntos no son isomorfos:

Lo posible es todo aquello que no es *necesariamente falso*, de modo que

[1] Es posible que Sócrates hubiera tenido el pelo rojo

significa

[2] Existe (al menos un) mundo posible *f* en el que la proposición «Hay algo que es Sócrates y que tiene el pelo rojo» es verdadera.

(...) [D]ebemos decidir si estamos hablando de posibilidad física o lógica. Tal hipótesis es *físicamente* posible si existe un mundo dentro del conjunto Φ [de los mundos físicamente posibles] donde hay un Sócrates pelirrojo. En caso contrario queda descartado desde el punto de vista físico, por más comunes que sean los Sócrates pelirrojos en los mundos lógicamente posibles pero físicamente imposibles.³⁷⁰

Hay pues que elegir con qué concepto de posibilidad queremos quedarnos. Recomendando que, dado que el mundo en el que tenemos que vivir es físico y no meramente lógico, optemos por la posibilidad física para hablar de libertad.

elemento líquido hacia arriba. De hecho, cuando una ballena queda varada no muere por asfixia, sino aplastada por su propio peso.

³⁷⁰ Dennett 2004: 88.

En cuanto a la necesidad cabe hacer la misma distinción. Sobre esa base, a lo largo de la historia de la filosofía han aparecido diferentes formas de concebir la necesidad. Así, por señalar tan sólo algunas de ellas,³⁷¹ Anselmo de Aosta considera que puede hablarse de algo que es necesario y contingente al mismo tiempo, aunque no en el mismo sentido; Tomás de Aquino piensa que hay que distinguir entre “necesidad absoluta” y “necesidad hipotética”; y Spinoza que si bien las esencias son contingentes, todo lo que es causado es necesario. Eso sí, siempre que no se trate de Dios, pues como “sabemos” por el argumento ontológico, en Dios esencia y existencia son uno y lo mismo. Pero lo que aquí nos interesa somos nosotros, los seres humanos y su mundo, y a este respecto hay que decir que éste es a la vez necesario y contingente: es necesario a posteriori, podría no haber sido pero es, y dado que es tiene una causa por la que es y por esa causa es necesario y es necesario que sea el que es; pero, por otro lado, no hay nada en su definición, en su esencia, que obligue a su existencia, y por ello es contingente. En este sentido hay que entender la divisa existencialista de que “la existencia precede a la esencia”. Y en este sentido hay que entender también las siguientes palabras de Anselmo:

Porque del hecho que se pone la existencia de una cosa, se dice que existe necesariamente; o si se pone su no existencia, se afirma su no existencia como necesaria, pero no porque la necesidad la fuerce o la impida ser o no. Cuando yo digo: «Si ella ha de existir, existirá necesariamente», la necesidad sigue aquí a la existencia de la cosa, pero no la precede. Lo mismo ocurre si digo: «Lo que ha de ser necesariamente será». Esta necesidad no quiere decir más que aquello que ha de existir no podrá al mismo tiempo no existir.³⁷²

Cuando algo ya existe, sería contradictorio decir que no existe. El hecho de la existencia convierte algo contingente en necesario, la posibilidad física hecha real entraña necesidad lógica en lo que a la existencia se refiere (pues lo

³⁷¹ He optado aquí por recoger argumentos de autores clásicos de la historia de la filosofía, pero no puedo dejar de mencionar el indispensable trabajo de Saul Kripke “Naming and Necessity” de 1972, en Davidson, D. & Harman, G. (eds) 1972, *Semantics of Natural Language*, Dordrecht: D. Reidel.

³⁷² Anselmo de Aosta 1953: 211.

“hecho real” y lo “existente” son lo mismo). Y algo así, como anuncié en la introducción histórica, defiende también Tomás de Aquino:

Una cosa puede ser necesaria, o absoluta o hipotéticamente. Es absolutamente necesaria por la correlación de los términos: bien porque el predicado está incluido en la definición del sujeto, como cuando decimos que es necesario que el hombre sea animal; o bien porque el sujeto pertenece a la razón del predicado, como cuando decimos que es necesario que el número sea par o impar. Y así, no es necesario que Sócrates esté sentado; es decir no es necesario absolutamente, pero puede decirse que es necesario hipotéticamente, pues en el supuesto de que esté sentado es necesario que esté sentado mientras permanece en esa posición.³⁷³

Lo que pretenden los libertaristas es que, una vez Sócrates está sentado, siga siendo posible que no lo esté, porque eso quiere decir que dos cosas son posibles simultáneamente o en exactamente las mismas circunstancias. Pero decir “Sócrates podría no haberse sentado” no puede querer decir que, una vez sentado Sócrates, ese hecho no era necesario hipotéticamente, esto es, físicamente, aunque es bien cierto que no lo es absolutamente, esto es, lógicamente. Pero esto último, insisto, no es sino señalar una obviedad, pues no quiere decir sino que no es autocontradictorio pensar que Sócrates podría estar de pie, lo cual es cierto, *siempre*. Hay cosas que son necesarias lógicamente y que no tiene sentido pensar que sean de otro modo que como son, y otras que son necesarias empíricamente y que tiene sentido pensar que sean de otro modo que como son, aunque sea imposible que lo sean *de hecho*. Que algo es materialmente contingente quiere decir que cabe imaginarlo siendo de otra forma que como efectivamente es sin caer en una contradicción, lo cual no quiere decir que efectivamente ese algo pueda ser de otra forma que como es, sólo que no es autocontradictorio verlo así. A modo de resumen de estas ideas, y como anticipo de unas consideraciones que voy a hacer acerca del azar, me voy a permitir transcribir gran parte del Escolio I a la Proposición XXXIII de la primera parte de la *Ética* de Spinoza:

³⁷³ Tomás de Aquino 1942: 304-305.

Se llama «necesaria» a una cosa, ya en razón de su esencia, ya en razón de su causa. [...] Además, por iguales motivos, se llama «imposible» a una cosa: o porque su esencia –o sea, su definición– implica contradicción, o porque no hay causa externa alguna determinada a producir tal cosa. Pero una cosa se llama «contingente» sólo con respecto a una deficiencia de nuestro conocimiento. En efecto, una cosa de cuya esencia ignoramos si implica contradicción, o de la que sabemos bien que no implica contradicción alguna, pero sin poder afirmar nada cierto de su existencia, porque se nos oculta el orden de las causas; tal cosa –digo– nunca puede aparecérsenos como necesaria, ni como imposible, y por eso la llamamos contingente o posible.³⁷⁴

Así, nada es contingente, la contingencia es una ilusión porque no hay nada casual y, en realidad, no existe el azar, como bien sabía Boecio:

Si por azar entendemos un acontecimiento fortuito, sin ningún nexo causal, hay que afirmar que el azar no existe y, de igual manera, que esta palabra (...) carece completamente de sentido. [...] [S]i un acontecimiento sucede sin causa lógicamente es como si saliera de la nada. Y si esto es imposible, también lo será la existencia del azar, según lo acabamos de definir. [...] Se habla de azar siempre que se hace algo con un fin determinado y por ciertas razones sucede algo distinto a lo previsto. (...) Se cree, entonces, que esto se ha debido al azar y que proviene de la nada. Pero tiene sus propias causas, cuya conjunción imprevista e inesperada da la impresión de que se ha producido algo casual.³⁷⁵

No deben definirse la contingencia y la posibilidad como opuestas a la necesidad, sino con respecto a la definición de aquello de que se trata, y con respecto a la posibilidad de predecirlo, respectivamente. Así lo hace Spinoza en la ética:

III. Llamo *contingentes* a las cosas singulares, en cuanto que, atendiendo a su sola esencia, no hallamos nada que afirme o excluya necesariamente su existencia.

IV. Llamo *posibles* a esas misma cosas singulares, en cuanto que, atendiendo a las causas en cuya virtud deben ser producidas, no sabemos si esas causas están determinadas a producirlas.³⁷⁶

Así, la contingencia hace referencia a la necesidad lógica que se deriva del hecho de que la definición de algo incluya lo que debe ser predicado de ello necesariamente. De lo que nos informa la contingencia de algo es acerca de cómo es, y no acerca de su posibilidad o imposibilidad de ser. Pero la

³⁷⁴ Spinoza 1988: 89-90.

³⁷⁵ Boecio 1999: 164.

posibilidad tampoco nos dice nada acerca de la necesidad de que algo sea, sino sobre nuestra capacidad para conocer ese algo y lo que éste podría producir. Las oposiciones necesario/posible y necesario/contingente no son correctas: las oposiciones reales son motivado/casual o a propósito/puramente efectivo. Esos son los términos que importan para el problema de la libertad, pues éste no es tanto un problema lógico como uno práctico en que hay que distinguir entre lo que sucede porque sí y lo que sucede porque hay una voluntad detrás. Si uno fuese capaz de penetrar de forma perfecta en los deseos y creencias de una persona, sería capaz de conocer sus motivos, el peso de cada uno de ellos y con ello de predecir sus acciones:

Así, si por necesidad entendiésemos la determinación cierta del hombre, que un perfecto conocimiento de todas las circunstancias de lo que sucede en el interior y el exterior del hombre, podría hacer prever a un espíritu perfecto, es seguro, que los pensamientos estando tan determinados como los movimientos, (...) todo acto libre sería necesario. Pero hay que distinguir lo necesario de lo contingente aunque determinado; y no solamente las verdades contingentes no son en absoluto necesarias, sino que incluso sus relaciones no son siempre de una necesidad absoluta, ya que hay que admitir, que existen diferencias en la manera de determinar entre las consecuencias que tienen lugar en materia necesaria y aquellas que tiene lugar en materia contingente. Las consecuencias Geométricas y Metafísicas necesitan, pero las consecuencias Físicas y Morales inclinan sin necesitar...³⁷⁷

Cualquiera que hubiese conocido el carácter de Edipo habría adivinado que iba a matar a Layo antes que dejarle pasar. Pero nadie es capaz de un conocimiento tan profundo, y por eso decimos que nuestras acciones no son necesarias a menos que, precisamente, seamos capaces de explicarlas en términos meramente causales, como sucede cuando, por ejemplo, explicamos el comportamiento de un esquizofrénico. Pero entonces se suspende el curso de las razones y se sustituye por el de las causas: “Existe el libre albedrío. No hay un ser dotado de razón falto de libertad. Todo aquel que por su misma naturaleza puede servirse de la razón tiene la facultad de poder juzgar y discernirlo todo. Por sí mismo distingue lo que ha de rechazar y lo que ha de

³⁷⁶ Spinoza 1988: 288.

³⁷⁷ Leibniz, *Nouveaux Essais sur l'Entendement Humain* pf. 13, en Leibniz 1959 : 253-254.

elegir.”³⁷⁸ Es cierto que las razones son causas, pero explicar mediante razones y explicar mediante causas tiene su contexto lingüístico apropiado,³⁷⁹ y el determinismo no atenta contra esa diferencia de contextos. Podría atentar contra él cierta forma de reduccionismo similar al conductismo... pero tal vez deberían ya los libertaristas, a estas alturas del debate, dejar de acusar a los compatibilistas de ser partidarios de Walden2.

“(…) [A]ún concediendo que la voluntad pertenezca a la esencia de Dios, no por ello deja de seguirse de su perfección que las cosas no han podido ser creadas por Dios de ninguna otra manera ni en ningún otro orden.”³⁸⁰ Si se adapta este texto de Spinoza a un vocabulario más neutral tenemos que: dado un principio, dado el mundo, éste es el que es y no podría ser otro, pero ello no implica que sea necesario, porque el que el principio fuera el que fue es contingente; no podría ser otro que el que fue, puesto que fue, pero no hay ninguna razón que haga necesario que fuera precisamente ese. El hecho es que el mundo es determinista y que en él empleamos el concepto de libertad con sentido, cuando menos con sentido pragmático. Tal vez, en lugar de preocuparnos por Walden2 o de buscar significados para “necesidad” y “contingencia” ajenos al mundo, deberíamos empezar a interesarnos por tratar de desentrañar cuál es ese sentido de “libertad” que nos permite atribuir responsabilidad en un mundo determinista. Un primer paso en este sentido podría ser el de ver cuáles son las razones de nuestro miedo a Walden2, de dónde vienen lo que Dennett ha llamado “los ogros” de la libertad. Tal vez una pequeña genealogía de esos miedos podría hacernos entender que el determinismo asustó en su día a los filósofos porque éstos esperaban del mundo más de lo que nosotros, afortunadamente, esperamos hoy, y que,

³⁷⁸ Boecio 1999: 166.

³⁷⁹ Confío en que no hará falta aquí profundizar una vez más en la teoría de la acción de Davidson: la causa de una acción intencional es una razón que la explica, está dentro del mismo orden ontológico que una causa natural, es también un suceso, pero no es susceptible de ser expresada nomológicamente.

puesto que hemos dejado de compartir sus anhelos, tal vez deberíamos empezar a desembarazarnos de sus problemas.

“Todos los prejuicios que intento indicar aquí dependen de uno solo, a saber: el hecho de que los hombres supongan, comúnmente, que todas las cosas de la naturaleza actúan, al igual que ellos mismos, por razón de un fin, e incluso tienen por cierto que Dios mismo dirige todas las cosas hacia un cierto fin, pues dicen que Dios ha hecho todas las cosas con vistas al hombre (...).”³⁸¹ En efecto, da la impresión de que el miedo a que todos seamos títeres es fruto del hecho de que se crea en la existencia de un titiritero mayor. En el helenismo surgió el concepto de libertad, pero los estoicos no tenían problema alguno con la providencia, y los epicúreos defendían una teoría materialista de la libertad. El problema lo crean los pensadores cristianos y se lo transmiten en herencia a los pensadores modernos. Son los cristianos quienes inventan la noción de “creación”, y esa creación cristiana *ex nihilo* convierte la providencia divina en dependencia absoluta del hombre con respecto a Dios, que ya no es sólo demiurgo sino causa primera entendida como “voluntad” primera. La incompatibilidad de la libertad no es con el determinismo, no es con la naturaleza, sino con la determinación *divina*, con la naturaleza *divina*. Como reza la obra de Anselmo, la concordia problemática es “la concordia de la presciencia, de la predestinación y gracia divina con el libre albedrío”. El problema de la libertad con la causalidad de la naturaleza surge cuando se presupone una voluntad que rige esa causalidad. Así, el problema de la libertad lo crea Agustín por derivarlo del problema del mal. Esto no quiere decir que el concepto de libertad no precise análisis y discusión, sino que el análisis y la discusión que se han llevado a cabo a lo largo de la historia de la filosofía

³⁸⁰ Spinoza 1988: 91.

³⁸¹ Ibidem: 96.

acerca del concepto de libertad están demasiado marcados por el desasosiego de Agustín ante el problema del mal. Así lo explica Schopenhauer:

Por una parte [Agustín] no quiere conceder a la libertad la amplitud que le reconoce Pelagio, porque, de ese modo, se anularían el pecado original, la necesidad de la salvación y la gratuidad de la gracia, y el hombre se justificaría por sus propias fuerzas y se haría digno de la salvación. [...] Por otro lado, sin embargo, tres motivos le mueven a defender la libertad de la voluntad: 1º Su oposición a los maniqueos (...). 2º Esa ilusión natural (...) en virtud de la cual el «yo puedo hacer lo que quiero» se considera como libertad de la voluntad (...). 3º La necesidad de poner de acuerdo la responsabilidad moral del hombre con la justicia de Dios.³⁸²

Este tercer motivo, fruto del rechazo a la idea de que Dios es autor del mal, se deja ver en la mayor parte de los autores libertaristas de la historia de la filosofía, en algunos de forma más o menos encubierta, como en Kant, en otros de forma totalmente abierta, como en Schelling. Pero si se elimina a Dios de nuestra mente, el problema de la libertad deja de ser tan desasosegante como lo era para Agustín, pues los males que provoca esta entelequia teológica sólo pueden combatirse con razones “que no son filosóficas, sino teológicas, es decir, que no gozan de validez absoluta”.³⁸³ La teología es una herramienta que no está en mi mano, y me gustaría creer que tampoco lo está en la de la mayor parte de los filósofos de nuestro tiempo, pero sin ella es imposible resolver el problema de la libertad tal y como éste aparece reflejado en la filosofía de los pensadores cristianos: “Pero ¿cómo representarnos que un ser, que en su existencia y en su esencia es, por completo, la obra de otro, se determine, sin embargo, radicalmente a sí mismo y pueda ser responsable por lo tanto, de su acción? [...] ¿Qué diríamos del relojero que se enfadara con su reloj porque éste no marchaba bien?”³⁸⁴ Si eliminamos al relojero desaparece el miedo a que seamos relojes. El problema, naturalmente, es que tantos siglos presuponiendo un relojero han acabado por conseguir que ahora veamos manos dispuestas a darnos cuerda por todas partes, aún cuando ya no creamos en el relojero.

³⁸² Schopenhauer 2000: 119-120.

³⁸³ Ibidem: 126.

“(…) [N]o hay ninguna contradicción entre la presciencia, a la que sigue la necesidad, y el libre albedrío, que está exento de ella, puesto que, de una parte, es necesario que sea lo que Dios conoce de antemano, y por otra, Dios sabe de antemano que hay futuros que están fuera de toda necesidad. […] Síguese, por consiguiente, que, sea que peques o que no peques, no habrá ninguna necesidad en uno u otro caso, porque Dios sabe de antemano que lo que será ocurrirá sin necesidad.”³⁸⁵ Así, Anselmo pretende que hay sucesos que no son necesarios, pero que sí son predecibles (por Dios), y sin embargo la realidad es exactamente al revés: *hay sucesos que son necesarios, pero impredecibles* (a menos que uno sea Dios). Precisamente, los problemas conceptuales serios en lo concerniente a la libertad surgen cuando se presupone un ser omnisciente y omnipotente. De hecho Descartes, en sus *Meditaciones metafísicas*, es capaz de elaborar la siguiente teoría compatibilista de la libertad, que ya cité más arriba:

Ya que [nuestra libertad] consiste solamente en que podemos hacer una cosa o no hacerla (...); o más bien solamente en que, para afirmar y negar, ir detrás o huir de las cosas que el entendimiento nos propone, actuamos de tal forma que no sentimos en absoluto que ninguna fuerza exterior nos constriña. Pues, para que sea libre, no es necesario que sea indiferente a elegir uno u otro de los dos contrarios; sino más bien, tanto más que me incline hacia uno (...).³⁸⁶

Pero claro, estas palabras se deben a que Descartes está hablando de la libertad humana, cuando prueba a hablar de la libertad divina y de cómo ésta y la humana encajan, entonces se ve sin más salida que recurrir al *liberum arbitrium indifferentiae*, esto es, a la idea de que tener libertad de elección es ser indiferente ante dos posibles acciones. Exactamente aquello que, en el caso de los seres humanos, describe como un defecto.

Los pensadores cristianos se han empeñado en que la libertad es compatible con la predestinación, la presciencia y la gracia divinas, pero

³⁸⁴ Ibidem.

³⁸⁵ Anselmo de Aosta 1953: 209.

³⁸⁶ Descartes 1996: 88-89.

incompatible con la necesidad, y que hay pues una predestinación que no determina: pero es justo al contrario. La tercera antinomia kantiana es un callejón sin salida en el que se enfrentan dos concepciones cristianas de la libertad, desde el materialismo monista no hay antinomia. La aparente contradicción que existe entre el determinismo y las posibilidades alternativas es una ilusión creada por la evolución filosófica de una contradicción real, a saber: “(...) cómo puede ser que Dios conozca de antemano todos los futuros y que, no obstante, no pequemos por necesidad. Porque decir que las cosas pueden suceder de otro modo distinto de como Dios las prevé, es intentar destruir la presciencia de Dios (...). [...] [P]or lo mismo que Dios había previsto que había de pecar, era de necesidad que sucediera lo que Dios había previsto que sucediera. ¿Cómo, pregunto yo ahora, puede darse la libertad de la voluntad donde tan evidente e inevitable es la necesidad?”³⁸⁷

Abandonemos la contradicción originaria y poco a poco seremos capaces de darnos cuenta de que hoy en día ya no hay contradicción alguna entre determinismo y libertad sino todo lo contrario. Así, el miedo del libertarista no es miedo a la determinación, sino a la predeterminación. Pues bien, el libertarista está de suerte: *no hay predeterminación*. Este hecho es importante porque, como ya se anticipó más arriba, puede hablarse de cierta transitividad de la responsabilidad si consideramos que la causa primera de toda la cadena causal que impulsa una acción es una causa *consciente*, lo cual viene a ser decir que esa causa primera es una *voluntad*. Agustín de Hipona tiene un problema con la libertad de los seres humanos porque su causa primera tiene una voluntad que determina necesariamente. No obstante, desde el materialismo, no hay tal problema porque la causa primera del mundo es contingente, o cuando menos inconsciente. No hay voluntad primera, sólo hay hecho puro, no hay acción sino suceso. Y nuestra voluntad no puede volcarse

³⁸⁷ Agustín 1971: 344 y 345.

transitivamente en ese suceso, porque éste no es voluntad sino simple causa, necesaria sí, pero no intencionada, una causa simplemente estúpida. Dicho de otro modo: nuestro futuro está determinado pero no está escrito en ningún lado. El oráculo de Delfos nunca tuvo ni idea de cuál había de ser el destino de Edipo. Para adivinar cuál iba a ser el destino de Edipo habría sido necesario, como mínimo: conocer la reacción de Layo ante la amenaza del oráculo, el comportamiento del encargado de asesinar a Edipo, el de los padres adoptivos de Edipo, la evolución del carácter de éste, el momento en el que iba a decidir abandonar Corinto, el momento en el que Layo iba a decidir ir en pos del oráculo de Delfos para pedirle consejo sobre la esfinge, el lugar en que padre e hijo se iban a encontrar y el estado anímico de ambos en ese momento. Sólo un tipo de ser podría ser capaz de conocer estos datos y por lo tanto de conocer la suerte de Edipo: Dios. Por eso lo único que hay incompatible con la libertad es un dios omnisciente y omnipotente. El determinismo admite perfectamente que las acciones humanas sean impredecibles, lo cual no quiere decir que conociendo bien a una persona no se pueda aventurar correctamente lo que haría en una situación determinada. Pero este punto es perfectamente coherente con la libertad, no lo es, en cambio la predeterminación, que es lo que hace que Dios conozca cualquiera de nuestros actos antes de que los hayamos realizado. Si Apolo, inspirador del oráculo de Delfos, fue capaz de predecir el futuro de Edipo no fue porque, dado el determinismo, fuese capaz de deducir de todos los elementos causales implicados lo que iba a pasar, sino porque, dado que él mismo lo había decidido *qua* Dios, ya sabía lo que iba a pasar. Los oráculos no predicen, desvelan. Los dioses no adivinan, saben. Los seres humanos sí predecimos y adivinamos, a veces, porque somos falibles. El que los seres humanos tengamos la capacidad de, dadas unas causas, deducir sus efectos, no conlleva que siempre vayamos a acertar. Y esto es perfectamente compatible con la libertad. Otra cosa sería que fuésemos capaces de elaborar leyes predictivas sobre el comportamiento humano

(cuestión esta de la que me ocuparé en II-3. y II-4.), pero no lo somos. Este extremo preocupa sobremanera al libertarista, y por eso éste postula el indeterminismo. Pero, para su tranquilidad, no es necesario que el indeterminismo se dé en las acciones de los seres humanos, tan sólo *en nuestra capacidad para conocer* la multiplicidad de causas que son su origen,³⁸⁸ porque ello es lo que nos permite funcionar como seres libres y responsables, y no como científicos prediciéndonos mutuamente el comportamiento (insisto en que este punto será más, incluso muy, ampliamente desarrollado en II-3. y II-4.).

En fin, he aquí otra hipótesis de la presentación corroborada, la 3), “es posible predecir comportamientos, sin embargo las acciones humanas poseen un buen grado de imprevisibilidad”.

I.3.2. La personalidad y su carácter: unidad del agente y sujeto de la libertad

Más arriba, cuando me ocupé de la crítica a las posibilidades alternativas, distinguí entre estrategias de coacción de las cuales el agente es consciente y estrategias de las cuales no lo es. Se vio que la coacción por sí sola no es capaz de explicar el que alguien haga A más bien que B, aunque es razón suficiente de que lo haga. Hay elementos coercitivos que son relevantes para el proceso de deliberación de un agente, y hay elementos coercitivos que no lo son. Para ser relevante en la elección de la acción, la coacción ha de ser conocida por el agente, porque de ese modo puede confrontarse con sus motivos, de lo contrario tan sólo puede hacerlo con sus acciones, esto es, después de la decisión. Esta necesidad de que el agente sea consciente de la coacción a la que es sometido para que ésta sea relevante para su elección, no

³⁸⁸ Por fin doy por cerrado otro tema que dejé abierto más atrás (véanse las notas 198 y 200): el de la diferencia que existe entre un hecho causado indeterminísticamente y mi incapacidad para

hace sino subrayar la importancia de la conciencia en lo que respecta al ejercicio de la libertad. Precisamente es la conciencia lo que convierte a las causas en motivos.

Si es cierto el determinismo, y la libertad no es sino causalidad, ¿qué es lo que nos permite hablar de libertad en unos casos y de ausencia de libertad en otros? Por lo visto hasta ahora se puede hablar de libertad si un agente, cuando actúa, controla su decisión y su acción, y si en circunstancias similares sería capaz de obrar de otro modo. Una forma de resumir estas instancias sería decir que *la libertad es una causalidad consciente*.³⁸⁹ Esta conciencia de la causalidad implica, como toda conciencia, reflexividad, esto es, no sólo que el agente actúe, sino también que se dé cuenta de que actúa (y que por lo tanto pueda evaluar y reconsiderar su acción), lo cual viene a ser una forma de autoconsciencia. Así, un ser libre sería una causa autoconsciente. Pero no sólo eso: “Estamos particularmente preocupados por nuestros propios motivos. Es muy importante para nosotros si los deseos por los que somos movidos a actuar del modo en que lo hacemos nos motivan porque queremos que sean efectivos a la hora de movernos o si nos mueven al margen de nosotros o incluso a pesar de nosotros.”³⁹⁰

La cuestión de los motivos de la acción está subsumida en esa parte de la libertad que en el párrafo anterior ha aparecido así: “el agente (...) controla su decisión y su acción”. Muchas cosas intervienen en la acción, pero puede decirse que, en general, y lamento no poder extenderme en esto más de lo necesario, los motivos de una acción se componen de deseos y razones, que a su vez, poniéndome de nuevo davidsoniano, son fruto de los deseos y creencias del agente. Así, actuar libremente vendría a ser actuar por motivos de

determinar la causa de un hecho, esto es, entre la indeterminación y la impredecibilidad.

³⁸⁹ Mi deuda con Spinoza es evidente.

³⁹⁰ Frankfurt 1987, en Frankfurt 1988: 163.

los cuales uno sería capaz de dar cuenta en términos de deseos y razones que uno reconocería como propios. Pero, una vez más, no basta con eso.

Considero que la libertad es no sólo una causa autoconsciente, sino *una causa autoconsciente con sentimiento de responsabilidad*. Esto implica lo siguiente, que al cuadro del agente con motivos compuestos por deseos y razones, hay que sumarle un conjunto de expectativas en que se basa la responsabilidad. Así, tendríamos que un agente libre es un ser autoconsciente con deseos y razones que motivan sus acciones, y con expectativas. Lamento no poder precisar aquí con más claridad de qué tipo de expectativas se trata, pues eso deberá ser estudiado de forma pormenorizada en la siguiente parte de este trabajo, pero tampoco es necesario para la cuestión que ahora nos ocupa, a saber: cuál es el sujeto de la libertad. Más adelante, tirando del hilo de las expectativas, vendré a decir que alguien responsable es alguien que vive en un mundo participante, de momento diré que es alguien que actúa de acuerdo a su *personalidad*, esto es, al conjunto de sus deseos, razones y expectativas. Voy a llamar aquí “personalidad” a un compuesto: el de aquello a lo que se le ha llamado tradicionalmente “carácter” (así lo hace Schopenhauer, por ejemplo) y a un conjunto ordenado de expectativas. Considero que hablar de personalidad posee dos ventajas: por un lado ésta puede ser entendida como *natura naturans* además de *natura naturata*; y por otro nos permite acercarnos a la psicología y a la psiquiatría, cosa ésta que, como intentaré mostrar en el último capítulo de este trabajo, puede arrojar mucha luz sobre el problema que nos ocupa. De momento sólo voy a tratar de la parte *natura naturans* de la personalidad, más adelante me ocuparé de la parte *natura naturata*.³⁹¹ Lo que nos concierne ahora tiene que ver con la parte individual de la personalidad, esto es, con el carácter; más adelante (en *III.4.1.*) distinguiré otra parte de la personalidad, una social,

³⁹¹ Concretamente en el apartado *III.4.1.* que he titulado *La personalidad y su virtud: las dos caras de la personalidad* y que viene a cerrar el problema que ahora estoy planteando en este *I.3.2. La*

que tiene que ver con la pertenencia a una comunidad moral.³⁹² Otra forma de distinguir ambas partes es decir que una es interna y la otra externa, lo cual se ajustaría a la distinción que hace John Martin Fischer entre teóricos internistas de la responsabilidad como Harry Frankfurt o Gerald Dworkin, y teóricos externistas como Peter F. Strawson o Robert Jay Wallace. Así, la parte individual o interna dependería del control que tenga el agente sobre sus motivos, y la parte social o externa del control que tenga el agente sobre sus expectativas y las de los demás. Quede esto dicho como anticipo, pues no puedo ir más allá de momento, en cualquier caso la teoría de la responsabilidad que voy a proponer, si bien insistirá más en la parte externa, puede considerarse mixta en el sentido de que reconozco la importancia del carácter como base de los motivos de las acciones. Ahora, lo que entiendo por carácter, como se verá en breve, está muy lejos de poder considerarse una lectura internista del mismo. Podría decirse más bien, que reconozco la importancia de la parte individual de la responsabilidad, pero no concedería que dicha parte individual se asienta exclusivamente en hechos internos de la psicología del agente. En cualquier caso, hasta que me sea posible dar cuenta de los factores externos que construyen la personalidad,³⁹³ habrá que conformarse con un primer sentido de “personalidad” como sujeto de la libertad, esto es, como el lugar en que se alojan los motivos de la acción: la personalidad como carácter. Sólo en este sentido la personalidad ya es suficientemente complicada de entender.

personalidad y su carácter: unidad del agente y sujeto de la libertad. Ambos apartados forman una unidad conceptual, pero no han podido verse reunidos físicamente en estas páginas.

³⁹² Insisto que hasta entonces no podrá considerarse que he logrado construir una teoría completa de la personalidad: el carácter y la virtud (los valores), lo personal y lo social, son partes igualmente importantes de la personalidad (o, como diré más adelante, son dos caras de la responsabilidad).

³⁹³ Véase la dos notas anteriores.

Los motivos no siempre forman un todo armónico; a veces, cuando el agente debe tomar una decisión, da la impresión de que hubiese dos agentes en vez de uno disputándose el derecho a elegir: “Cuando alguien se identifica con uno más bien que con otro de sus deseos, el resultado no es necesariamente la eliminación del conflicto entre esos deseos, o la reducción de su severidad, sino la alteración de su naturaleza.”³⁹⁴ A menudo, y en el fondo son estos conflictos los que le dan sabor al ejercicio de la libertad, hay que elegir entre motivos contradictorios que son, a pesar de ello, fruto igualmente de nuestros deseos y razones. Este hecho ha llevado a numerosos autores a esforzarse por elaborar teorías sobre el agente libre que, en vez de afanarse en explicar el hecho de las posibilidades alternativas, salven la identidad del agente dando una explicación para estos conflictos de motivos en términos de identificación del agente con unos motivos y no con otros. Y una vez más un autor a destacar dentro de esta corriente es Harry Frankfurt.

Como ya ha sido dicho Frankfurt pertenece, además de Charles Taylor, Gerald Dworkin y Michael Bratman, entre otros, a un grupo de filósofos que basan la libertad de las acciones en el hecho de que éstas dependan de, en palabras de Susan Wolf, “el verdadero yo” del agente. Todas estas teorías, con matices³⁹⁵, describen jerárquicamente la personalidad del agente, de tal forma que todos nuestros motivos acaban remitiendo a un nivel básico de deseos que componen nuestro carácter. Así, Frankfurt propone un “modelo que implica niveles de reflexividad o de autoconsciencia. De acuerdo con este esquema, en el nivel más bajo están los deseos de primer orden de realizar una u otra acción. Cualquiera de estos deseos de primer orden que efectivamente conduzca a la acción es, en virtud de esa efectividad, designado por la voluntad del individuo a quien pertenece ese deseo. Además, la gente se caracteriza por

³⁹⁴ Ibidem: 172.

³⁹⁵ No me detendré en los matices. A este respecto consúltense Frankfurt 1971, Taylor 1976, Dworkin 1988 y Bratman 1999. También Watson 1975.

tener deseos de segundo orden acerca de los deseos de primer orden que quieren, y por tener voliciones de segundo orden acerca de qué deseos de primer orden quieren que sea su voluntad. También puede haber deseos y voliciones de niveles más altos.”³⁹⁶ Así, nuestro carácter estaría compuesto por deseos de segundo orden, que nos permitirían decidir, dado un conflicto concreto entre dos deseos de primer orden concretos, cual de esos deseos o motivos queremos que cause nuestra acción. Y así es como nuestro carácter funciona como *natura naturans*, porque al elegir qué deseo queremos que sea nuestro motivo, estamos decidiendo qué respuesta a un conflicto dado es *nuestra* respuesta, y por lo tanto quienes somos y cuál es nuestro carácter: “(...) la persona, al tomar una decisión por la que se identifica con un deseo, *se constituye a sí misma*.”³⁹⁷ Ciertamente esta idea recuerda al concepto de “acciones auto-formativas” de Kane, sin embargo hay notables diferencias entre ambas nociones. La más importante: Frankfurt ni tan siquiera se plantea que tenga que tratarse de decisiones indeterminadas, es más, si las personas funcionan como *natura naturans* es porque precisamente son como la sustancia de Spinoza, un sistema plenamente determinista. Tal y como quiere también Schopenhauer: “(...) la voluntad humana constituye el auténtico yo del hombre, el núcleo verdadero de su ser (...). De aquí que preguntarle si también podría querer de otra manera a como quiere vale tanto como preguntarle si podría ser otro del que es; y esto sí que no sabe.”³⁹⁸

El problema al que se enfrenta Frankfurt, y que intentará resolver infructuosamente con la idea de decisiones tomadas “con el corazón en la mano”³⁹⁹ (*wholehearted*), es el siguiente: es posible que un agente tenga deseos de segundo orden que estén en conflicto, y para resolver ese conflicto sería

³⁹⁶ Frankfurt 1987, en Frankfurt 1988: 164.

³⁹⁷ Ibidem: 170.

³⁹⁸ Schopenhauer 2000: 66.

³⁹⁹ La traducción no es exacta, pero creo que así es más significativa en castellano de lo que lo sería la traducción literal.

necesario acudir a deseos de orden superior, y de darse un conflicto también entre estos, habría que acudir a otros de orden superior y así, ¿hasta dónde? Ya Leibniz había hurgado en esta herida varios siglos atrás: “(...) haciendo que la acción de querer dependa de la voluntad, tiene que haber otra voluntad o facultad de querer anterior para determinar los actos de esa voluntad, y otra para determinar ésta, y así hasta el infinito; ya que se pare uno donde se pare, las acciones de la última voluntad no sabrían ser libres.”⁴⁰⁰ El problema de la teoría jerárquica de Frankfurt es pues de regresión al infinito, una regresión que hace imposible que se atribuya responsabilidad por una acción, pues la cadena causal de deseos de diferentes órdenes debería acabar llevando hasta los genes y los factores ambientales que formaron los primeros deseos del agente. Mi idea es que este problema no demuestra que la teoría de Frankfurt o variantes similares como la de Taylor estén equivocadas, tan sólo que son incompletas porque se centran en un único aspecto de la personalidad, el carácter, y olvidan el aspecto social. Pero que el cuadro no esté completo no significa que sea malo y que haya que mejorar lo ya pintado, tan sólo que hace falta ir a comprar más pintura y terminarlo.

A Frankfurt le ocurre, al fin, lo mismo que a los libertaristas: que necesita pensar que el sujeto es autoconstituido, que hay una barrera entre lo que los factores externos hacen de uno y lo que uno mismo hace de sí mismo. Pero me temo que esa barrera no existe, al menos tal y como le gustaría a Frankfurt que existiera. Dado el determinismo, si el carácter es la base de nuestra libertad, y los deseos de segundo orden la causa de nuestras acciones libres, hay que pensar que ambas cosas son también efectos:

[D]el mismo modo que es imposible que una bola de billar se mueva sin que reciba antes un golpe, así también es imposible que un hombre se levante de la silla sin un motivo que le empuje o lleve; con lo que su levantarse es algo tan necesario e inevitable como el rodar de la bola después del choque. [...] Porque el hombre, como cualquier otro objeto de la experiencia, es un fenómeno que se desarrolla en el tiempo y en el espacio, y como la ley de la causalidad se

⁴⁰⁰ Leibniz, *Nouveaux Essais sur l'Entendement Humain* pf. 22, en Leibniz 1959 : 255.

aplica *a priori* a todos ellos, lo que quiere decir sin excepción, también él tendrá que estar sometido a ella.⁴⁰¹

Ahora bien, que los agentes no sean autoconstituidos, no quiere decir que no se autoconstituyan mediante sus acciones. En último término el carácter es fruto de nuestros genes y de la educación, y como tal está ya dado pero, como bien señala Frankfurt, es también capaz de autoconstituirse recursivamente por el carácter reflexivo de la conciencia. En este sentido decía yo más arriba que es la conciencia lo que convierte a las causas en motivos. El carácter, donde se alojan los deseos de segundo orden, es un resultado que *dejamos de considerar efecto y por convención*, no por un tipo especial de acción mediante la cual el agente se funda a sí mismo, *empezamos a considerar causa* a partir de cierta edad. El carácter es un conjunto reorganizado de efectos que se convierten en motivos fruto de un todo organizado, jerarquizado y unificado. Cuando el carácter se ha constituido porque los demás, la sociedad, consideran que así es, las causas que lo determinaron se hacen irrelevantes. Para explicar sus acciones basta con las razones que dé el agente, razones que son motivos/deseos de primer orden, basados en principios/deseos de segundo orden. Sólo cuando estas razones son incomprensibles se hacen relevantes las causas que llevaron al agente a actuar de forma determinada desde un pasado remoto. Se trata de casos en los que hay trastornos de la personalidad. Existen personas cuya personalidad no se ha llegado a formar correctamente como un yo unificado y con sentido, y debido a ello algunas de sus acciones tienen una explicación puramente causal y no racional, por lo que para entenderlas es necesario buscar una causa última de las mismas en toda la cadena causal que llevó hasta ellas.⁴⁰²

La crítica que puede hacerse a estas ideas es obvia: ¿Dónde está ese yo?, ¿Cuándo puede decirse que ya tenemos un carácter constituido y dejamos de

⁴⁰¹ Schopenhauer 2000: 93.

⁴⁰² De todo esto me ocuparé de forma detallada en III-4.

ser efecto para convertirnos en causa? Me temo que no soy capaz de responder a estas preguntas, ni creo que nadie lo sea. Sin embargo, me da la impresión de que hacerse estas pregunta es como tratar de averiguar, como diría Dennett, cuál fue el primer mamífero de la historia, esto es, preguntarse por la esencia de cada uno. Pero esto no es necesario, cuando hablamos de acciones fruto del carácter de cada uno, o cuando damos el salto y empezamos a considerar a un niño autor de sus acciones, no solemos requerir exactitud, simplemente sabemos hacerlo, somos competentes en ello, conocemos cómo jugar ese juego. En el caso de los mamíferos, forma parte del ser mamífero el ser descendiente de un mamífero, pero el caso es que en algún momento de la evolución debió surgir un mamífero de algo que no lo era, así pues los mamíferos no existen. Pero es un hecho que los mamíferos existen y que somos capaces de distinguirlos, aún cuando no seamos capaces de decir exactamente en qué momento de la evolución algo dejó de ser un reptil y se convirtió en mamífero:

[S]i nos remontamos lo bastante en el árbol de familia de cualquier mamífero, llegaremos finalmente a los terápsidos, una extraña y extinta especie puente entre los reptiles y los mamíferos. Se produjo una transición gradual de los reptiles puros a los mamíferos puros, y el espacio intermedio fue cubierto por numerosos intermediarios de difícil clasificación. [...] ¿Podemos identificar algún mamífero, el Mamífero Primordial, que no tuviera a un mamífero por madre (...)? [...] Deberíamos poner coto a nuestro afán de marcar líneas divisorias. No necesitamos tales líneas. Podemos vivir con el hecho escasamente misterioso y sorprendente de que, ya ves, se produjeron una serie de cambios graduales que se fueron acumulando a lo largo de millones de años y al final dieron como resultado mamíferos puros.⁴⁰³

Yo diría más bien que “dieron como resultado *lo que decidimos considerar* mamíferos puros”, pero en cualquier caso preguntarse por el momento en el que el carácter pasa de ser constituido a constituyente viene a ser preguntarse por el “Mamífero Primordial”. Conocer con exactitud ese momento es tan poco necesario como conocer con exactitud al primer animal que pueda ser considerado un mamífero puro. En nuestras infancias se “produjeron una serie de cambios graduales que se fueron acumulando” y al final dieron como

resultado un carácter. Con esto es suficiente. El carácter es como una playa, no se sabe muy bien dónde acaba la arena y dónde empieza el mar, pero sí es fácil saber dónde hay que poner la toalla para no mojarse.

La libertad es pues una causalidad autoconsciente que se hace presente en el sentimiento de responsabilidad: “Si ahora, como consecuencia de lo que llevamos expuesto, eliminamos por completo la libertad de las acciones humanas reconociéndolas como sometidas absolutamente a la necesidad más rigurosa, habremos llegado al punto desde el cual podemos comprender la *verdadera libertad moral*, que es de un tipo superior. Existe otro dato de la conciencia (...). Se trata del sentimiento completamente claro y seguro, de la *responsabilidad* por aquello que hacemos, de la *imputabilidad* de nuestras acciones, lo que descansa sobre la certeza incommovible de que somos nosotros mismos los *autores de nuestros actos*.”⁴⁰⁴ Ese autor de sus actos, ese sujeto de la causalidad consciente, es el carácter, que determina las acciones del agente, siendo él mismo el producto del determinismo genético y ambiental. Una multitud de causas confluyen en un sujeto hasta reorganizarse en un todo suficientemente coherente y estructurado de creencias y deseos que constituyen un carácter y que mueven a la acción. Ese carácter *no* es pues autoconstituido (*ergo* no es *causa sui*), pero *sí* autoconstituyente, un todo independiente con deseos y motivos achacables a sí mismo y no a la multiplicidad incontable de causas que llevó a crear ese carácter (como he dicho, en esto hay excepciones, porque no todo carácter puede ser considerado “bien terminado”, y en esos casos hay que buscar la responsabilidad⁴⁰⁵ por algunas de las acciones de ese carácter en aquellos o aquello que lo mal terminó). Como defiende Spinoza, “voluntad y entendimiento son uno y lo mismo”, esto es, ser libre es ser autoconsciente de

⁴⁰³ Dennett 2004: 150.

⁴⁰⁴ Schopenhauer 2000: 151.

⁴⁰⁵ Es importante señalar lo siguiente: el carácter puede ser la base de la libertad de la voluntad, pero no de la responsabilidad. Ésta requiere de las dos caras de la personalidad: la individual y la

la determinación de nuestra personalidad, a través de sus motivos, sobre nuestras acciones. En la deliberación hacemos hipótesis sobre cual será el futuro dependiendo de nuestra acción. Precisamente actuamos porque sabemos que sólo un futuro es posible, y que éste depende absolutamente de nuestra elección, está completamente determinado por ella. Y por ello las acciones nos dicen quienes somos, porque nos desvelan nuestro carácter, algo que sólo nos es conocido plenamente *a posteriori*, y por ello los seres humanos somos en gran medida impredecibles (lo cual confunde al libertarista con el que estemos indeterminados). Si Edipo no hubiese matado a su padre y no se hubiese acostado con su madre, no habría sido Edipo. Así, Edipo no podía no ser Edipo. Pero, como ya dije antes, el oráculo no lo sabía.

social. Me temo que habrá que esperar hasta el último capítulo de este trabajo para encontrar la prueba de esta afirmación. Véanse las notas 391 y 392.

Recapitulación

En esta primera parte se han estudiado los fundamentos metafísicos de la libertad.

En un primer momento, al hilo del debate entre compatibilistas e incompatibilistas, se señalaron dos requisitos indispensables de la libertad de la voluntad: el de las responsabilidades alternativas y el del control del agente. Se apuntó que las teorías compatibilistas suelen tener problemas para cumplir el primer requisito y las incompatibilistas el segundo. A continuación se demostró que las diferentes versiones del incompatibilismo actualmente vigentes son incapaces de dar cuenta del control del agente sin postular conceptos metafísicos como la *causa sui*, el *clinamen* y el dualismo de sustancias. Se juzgó que dichos postulados no constituyen una solución al problema y se concluyó que es mejor optar por una visión compatibilista de la libertad. Esta conclusión lleva consigo la verdad de la hipótesis 1) de la presentación del trabajo: “Las acciones humanas están determinadas, determinadas por los propios seres humanos.”

Tras esta parte crítica se dio paso a una positiva de elaboración de una teoría compatibilista que cumpliera el requisito de las posibilidades alternativas. Para ello se llevó a cabo un estudio fenomenológico de la expresión “podría haber obrado de otro modo” que acabó por mostrar que “una o más personas pueden actuar de forma diferente en situaciones muy parecidas”, lo cual constituye la verificación de la hipótesis 2) de la presentación.

Por último, se trató de trabajar sobre el sujeto de la libertad, sobre la forma en que éste, a pesar de estar determinado, es capaz de ser libre y dueño de sí mismo. Dicho esfuerzo permitió falsar la idea de que el determinismo y la libertad son incompatibles en tanto dicha idea se asienta en una mala

comprensión de los conceptos de posibilidad y necesidad. También, se llegó a la conclusión de que el carácter de la persona es la base de la libertad de su voluntad, y que el determinismo no implica que las personas no puedan considerarse entidades autoconstituyentes a pesar de no ser autoconstituidas, ya que, aunque “es posible predecir comportamientos, sin embargo las acciones humanas poseen un buen grado de imprevisibilidad”, y esto conlleva que podemos juzgar las acciones de un agente como suyas propias y no como fruto de la causalidad universal. Esto último responde a la hipótesis 3) de la presentación.

En esta parte han quedado muchos hilos sueltos que voy a tratar de recoger en las dos partes siguientes. En ellas dejaré de lado los aspectos más metafísicos de la cuestión y pasaré a analizar la libertad como un hecho en ejercicio, concretamente como el ejercicio de la responsabilidad moral. Retomaré el tema del incompatibilismo, pero en esta ocasión no desde el punto de vista de la voluntad, sino de la responsabilidad. Haré pues una nueva defensa del compatibilismo mediante argumentos más prácticos y menos especulativos. Pero sobre todo trabajaré para construir un todo coherente que dé cuenta de aquello en que consiste ser responsable moralmente.

II- LA RESPONSABILIDAD MORAL

Aunque abriré esta segunda parte del trabajo recordando las intuiciones de Hume acerca de los sentimientos morales, enseguida me centraré en la filosofía del siglo XX. Voy a concentrarme en el análisis de la responsabilidad moral *como hecho* y voy a llevar a término esta labor utilizando del artículo de P. F. Strawson “Freedom and Resentment” como principal herramienta. También llevaré a cabo un análisis de los argumentos que utiliza Strawson en su defensa del compatibilismo, lo cual me permitirá hablar de dos estrategias para refutar los argumentos del incompatibilista en contra de la responsabilidad: una naturalista y otra racionalista. Aquí investigaré cuáles son las consecuencias de seguir cada una de estas dos estrategias y me preguntaré hasta qué punto son ambas conciliables. Mi idea es que ambas estrategias no pueden ser desarrolladas conjuntamente, al menos tal y como lo hace Strawson en su artículo. Dicha unión de fuerzas sólo es posible si se tiene en cuenta que desde cada una de estas estrategias sólo se puede rebatir una parte diferente del incompatibilismo. Así, la estrategia naturalista nos permite entender por qué, hoy por hoy, nos es imposible comportarnos como genuinos incompatibilistas, pero no nos permite afirmar que sería irracional intentar hacerlo. Por el contrario, la estrategia racionalista no nos permite fundamentar con fuerza nuestra autopercepción como seres intrínsecamente responsables, tal y como nos muestra la estrategia naturalista, pero sí nos permite afirmar que, dado el hecho de la libertad y la responsabilidad, entonces lo más racional es siguiendo vida a ese hecho. Como digo ambas estrategias se complementan, pero eso no las hace ni igualmente válidas ni igualmente necesarias, aquí apostaré sobre todo por la estrategia racionalista trabajada desde un punto de vista pragmatista.

Por ello, tras un breve repaso a algunos puntos de la teoría moral de David Hume, en un primer momento (II-1.) me ocuparé de analizar de forma pormenorizada “Freedom and Resentment” destacando sus ideas más novedosas y las más relevantes para el debate que nos ocupa. Daré cuenta de los principales argumentos que elaboró Strawson a favor de su teoría de la responsabilidad moral y de los principales conceptos filosóficos que manejó en su artículo.

En segundo lugar (II-2.) me interesaré por averiguar en qué consiste el naturalismo de Strawson y en qué su racionalismo. Veré cuáles son los problemas a los que tienen que enfrentarse cada una de estas estrategias y de qué modo podrían ambas, si posible, encajar tal y como pretendía Strawson.

En tercer lugar (II-3.) investigaré si se puede realizar alguna interpretación del naturalismo presente en el “Freedom and Resentment”, que haga a esa estrategia compatible con ciertos aspectos del racionalismo y capaz de soportar acusaciones que algunos autores han vertido contra ella por su carácter esencialista.

Por último (II-4.) mostraré las virtudes de la estrategia racionalista y la fuerza de sus argumentos contra la crítica incompatibilista. Hecho esto acabaré esbozando una teoría de la responsabilidad moral en una nueva versión pragmatista racional del normativismo strawsoniano.

Este todo debería permitirme confirmar la hipótesis 4) de este trabajo: “Tenemos expectativas y emitimos juicios de responsabilidad sobre las acciones ajenas y propias basándonos en dichas expectativas.”

II-1. El día a día de la responsabilidad moral

Es, pues, en estos sentimientos, y no en el descubrimiento de relaciones de ningún tipo, en lo que consisten todas las determinaciones morales. Antes de que podamos pretender llegar a alguna decisión de

esta clase, todo debe ser conocido y confirmado, por lo que al objeto o acción se refiere. Sólo nos queda, entonces, experimentar por nuestra parte un sentimiento de censura o de aprobación; y, a partir de ahí, nos pronunciamos acerca de si la acción es criminal o virtuosa.⁴⁰⁶

En su *Investigación sobre los principios de la moral*, David Hume relega el estudio del sentimiento moral a un apéndice al final del libro. La razón de este aislamiento impuesto a la que es la piedra miliar de su sistema moral sería que ocuparse de la misma supone un ejercicio de metafísica de la moral y no de filosofía moral propiamente dicha: “(...) decidimos posponer esa cuestión para no enredarnos en intrincadas especulaciones impropias de los discursos morales (...)”⁴⁰⁷ Esto es así precisamente porque, según Hume, del sentimiento moral es de lo que dependen los principios generales de la moral. Así, recorriendo un camino filosófico que siglos más tarde seguirá P. Strawson, Hume comienza analizando un criterio utilitarista como “fundamento principal de la alabanza moral”⁴⁰⁸ y lo presenta como paradigma de una concepción racionalista de la moral, pues sería la razón la que podría “instruirnos acerca de la tendencia de cualidades y acciones, y señalarnos sus consecuencias beneficiosas para la sociedad y para quien las posee”.⁴⁰⁹ Hume no llegará a desarrollar la cuestión tan sutilmente como Strawson, no obstante el giro de tuerca de “Freedom and Resentment” está ya en la *Investigación sobre los principios de la moral*: “Pero aunque la razón, cuando se ve cumplidamente asistida y mejorada, sea suficiente para instruirnos acerca de si las tendencias de las cualidades de las acciones son perniciosas o son útiles, no es por sí sola suficiente para producir ninguna censura o aprobación moral. La utilidad es sólo una tendencia hacia un cierto fin; y si el fin nos resultara totalmente indiferente, habríamos de sentir la misma indiferencia hacia los medios. Se requiere, pues, que un *sentimiento* se manifieste, a fin de dar preferencia a las

⁴⁰⁶ Hume 1993: 178.

⁴⁰⁷ Ibidem: 171.

⁴⁰⁸ Ibidem.

tendencias útiles sobre las perniciosas.”⁴¹⁰ El argumento que está esgrimiendo aquí Hume en contra del utilitarismo es una versión de su denuncia de la falacia naturalista: la razón sólo nos da acceso a los hechos, son los sentimientos los que nos permiten hacer juicios de valor. Pero lo que más interesa aquí no es esa estrategia humeana antirracionalista de sobra conocida, sino el hecho de que el filósofo escocés focalice la discusión en el problema de la responsabilidad moral, del cómo saber cuando podemos censurar o aprobar moralmente la conducta de alguien. Este hecho no se tiene siempre en cuenta cuando se habla del sistema moral humeano, y es lo que hace que se entienda su filosofía moral como un emotivismo moral desnudo de razones. Nada más lejos de la realidad: el sentimiento moral sólo puede sobrevenir cuando se conocen, mediante la razón por supuesto, todas las circunstancias del caso que se juzga. Ciertamente, según Hume es el sentimiento lo que dota de carga moral a nuestro juicio, que de lo contrario sería un mero juicio de hecho. Pero el citado sentimiento depende de la deliberación, del análisis racional del caso: “Se supone que todas las circunstancias del caso están ante nosotros antes de que podamos formular algún juicio de censura o aprobación. Y si alguna circunstancia material nos es todavía desconocida o dudosa, debemos primero emplear nuestra capacidad de investigación y nuestras facultades intelectuales en asegurarnos respecto a ella; y debemos, durante ese tiempo, suspender toda decisión o sentimiento moral.”⁴¹¹ Si Hume hubiese sido capaz de ser más preciso acerca de las circunstancias que nos llevan de vez en cuando a suspender nuestro sentimiento moral le habría tenido que dedicar todo lo que resta de trabajo a él; no obstante fue otro gran filósofo el que lo logró.

En “Freedom and Resentment” Strawson recoge el testigo de Hume e intenta naturalizar la cuestión metafísica de la libertad. Strawson sostiene que

⁴⁰⁹ Ibidem.

⁴¹⁰ Ibidem: 172-173. El subrayado es mío.

⁴¹¹ Ibidem: 177.

no es necesario contar con una definición clara de lo que es el determinismo para dar respuesta a la pregunta de si somos o no seres moralmente responsables. De hecho, el que la discusión acerca de la responsabilidad se hubiese enfangado hasta llegar a un callejón sin salida, como parecía ser el caso a principios de los años sesenta, sería, según Strawson, resultado del empecinamiento de muchos filósofos en dar una respuesta metafísica a un problema de carácter práctico como lo es el de la responsabilidad moral. La idea de Strawson es que no es posible dar respuesta a una pregunta general sobre el alcance del determinismo con respecto a lo seres humanos desde la filosofía, y por ello cabría esperar que un debate filosófico planteado en esos términos hubiese de desembocar siempre en una aporía. En cierto modo A. J. Ayer, en su artículo “Free Will and Rationality”⁴¹², refuerza esta idea cuando dice que las tesis deterministas son de facto insuficientes para hacernos cambiar nuestros hábitos morales porque no contamos con una descripción científica suficientemente detallada acerca de cómo actúa el determinismo en el comportamiento moral humano; o lo que es lo mismo, es la ciencia la que, suponiendo que algo así sea posible, podrá suministrarnos una explicación correcta del funcionamiento determinista del ser humano, y no la filosofía.

Sin embargo, tampoco cabe afirmar que pueda prescindirse de cierta idea de lo que es el determinismo para abordar el problema de la responsabilidad, lo cual no quiere decir que la solución a dicho problema vaya a venir de la mano de una definición clara de lo que el determinismo es. Lo que sí hay que tener claro, recuerda Strawson, es que existe una definición del determinismo tal que según ella éste es incompatible con la responsabilidad. El error no es pues tratar de definir lo que el determinismo es, sino intentar dar respuesta mediante ello a un problema práctico, a saber, el de si nuestras prácticas sociales de atribución de responsabilidad están o no justificadas. Ese

⁴¹² En Van Stratten 1980: 1-13.

error lo comete no sólo aquel que piensa que dichas prácticas deberían ser invalidadas por el determinismo, sino también cierta clase de compatibilista, el que da una definición del determinismo tal que piensa que justifica la permanencia de nuestras prácticas de imputación de responsabilidad. Cuando aquel que pretende mostrar que dichas prácticas están justificadas abandona el terreno de las justificaciones morales y se sumerge en el debate metafísico al que le insta el incompatibilista, pierde la razón porque acaba recurriendo a argumentos que: o bien podrían ser útiles para criticar la postura metafísica del incompatibilista, o bien podrían ser útiles para justificar nuestras prácticas morales, pero no ambas cosas.

Por ello, Strawson dedica gran parte de su artículo no sólo a criticar la postura incompatibilista, sino también a criticar algunas formas de defender nuestras prácticas morales dentro del debate acerca del determinismo, en concreto la estrategia utilitarista. Strawson considera que, al entrar en el juego del incompatibilista, en lo que respecta a investigar si nuestras prácticas de atribución de responsabilidad están o no justificadas, si bien el compatibilista está en la buena vía, equivoca sus argumentos del mismo modo en que lo hace el incompatibilista.

En palabras de Strawson la postura del compatibilista podría resumirse mediante los siguientes principios:

- (1) “Los hechos tal y como los conocemos no muestran que el determinismo sea falso.”⁴¹³
- (2) “Los hechos tal y como los conocemos aportan una base adecuada para los conceptos y las prácticas que el pesimista [léase incompatibilista] siente puestos en peligro por la posible verdad del determinismo.”⁴¹⁴

⁴¹³ Strawson 1962, en Fischer & Ravizza 1993: 46.

⁴¹⁴ Ibidem.

Strawson considera que así leída la postura del compatibilista es correcta, pero que sin embargo no da respuesta a los principales problemas que plantea un debate sobre aquello en que consiste ser un agente libre y por lo tanto responsable. En efecto, ¿por qué habría de justificar esta concepción de la libertad una práctica tal como la de la atribución de responsabilidades morales?

Prácticas como ésta, corrientes en nuestras sociedades, son las que parecen estar en juego si no conseguimos dotarnos de una teoría racional del agente responsable que las justifique, y constituyen, según Strawson, el verdadero problema a resolver. Al margen del problema metafísico en sí del determinismo y la libertad lo que le interesa a Strawson son las consecuencias prácticas de dicho dilema y, como se dijo más arriba, para estudiarlas una investigación en el plano metafísico resulta inadecuada. Por esta razón, antes de que Strawson publicase su artículo lo normal era que el compatibilista recurriese a un nuevo argumento de corte utilitarista para justificar la permanencia de nuestras prácticas de atribución de responsabilidad. Así, el argumento que acababa esgrimiendo el compatibilista para justificar la práctica del castigo u otras similares era defender que dichas prácticas son capaces de “regular el comportamiento humano de forma socialmente deseable”.⁴¹⁵ Como se vio hace un momento, este argumento en todo caso podría ser válido para justificar nuestras prácticas morales de atribución de responsabilidad, pero no lo es en absoluto para criticar la postura del incompatibilista, pues no constituye ni mucho menos una refutación de su tesis según la cual si el determinismo es cierto entonces *no es racional* concebir a los seres humanos como agentes libres y responsables. De hecho un incompatibilista podría cínicamente sostener el argumento utilitarista desde su propia postura: si se

⁴¹⁵ Ibidem.

trata de regular el comportamiento humano de forma socialmente positiva el debate acerca de la libertad es irrelevante.

Así pues, según Strawson queda una laguna en la argumentación del compatibilista ahí donde deberían unirse su postura, tal y como ha sido presentada más arriba, con nuestras prácticas sociales comunes de atribución de responsabilidad. En su artículo Strawson se ocupa sobre todo de llenar esa laguna y descubre que aquello que lo hace, y que por lo tanto funciona como justificación de nuestras prácticas morales de atribución de responsabilidad, son unas emociones, concretamente las actitudes reactivas.

II.1.1. La actitud participante y la actitud objetiva

El señalar la importancia de las actitudes reactivas fue la gran aportación de Strawson al debate acerca de la responsabilidad moral. Para Strawson tanto el incompatibilista como el compatibilista clásico no hacen sino “sobreintelectualizar” los hechos. La forma adecuada de estudiar el concepto de responsabilidad moral es acudir directamente a estudiar el *hecho* de que normalmente funcionamos moralmente haciendo atribuciones de responsabilidad, con respecto a los demás y con respecto a nosotros mismos. Se debería pues soslayar la discusión metafísica y darle la vuelta al problema. El filósofo debe preguntarse primero qué es lo que hacemos cuando realizamos atribuciones de responsabilidad, estudiar en qué consisten nuestras prácticas de rechazo y aprobación moral. Sólo en segundo lugar cabe preguntarse qué postura metafísica da mejor cuenta de estas prácticas y si desde ese punto de vista están éstas o no justificadas.

Bien, al preguntarse acerca del hecho de que normalmente funcionamos moralmente haciendo atribuciones de responsabilidad, Strawson descubre que cuando hacemos una imputación de responsabilidad ésta suele ir

de la mano de ciertas emociones a las que Strawson decide llamar *actitudes reactivas*. Se trata de emociones que llevan siempre consigo un juicio sobre intenciones (lo cual explica que constituyan atribuciones de responsabilidad) y que se dan en las personas cuando éstas se relacionan entre sí. Estas actitudes surgen en nosotros como reacción a una determinada acción llevada a cabo por un agente siempre que consideramos que dicho agente llevó a cabo su acción intencionadamente, esto es: las actitudes reactivas surgen cuando consideramos que alguien es responsable de un bien o un mal que nos ha sido hecho voluntariamente. Según Strawson si se trata de un bien sentimos actitudes de carácter positivo tales como la gratitud, la admiración o incluso el cariño. Si por el contrario se trata de un mal sentimos actitudes del rango del resentimiento, el odio o la decepción. Es importante insistir en que dichas actitudes sólo pueden tener lugar cuando aquel que las siente lleva a cabo un juicio sobre las intenciones de un agente (el sentir las constituye en sí mismo ese juicio): difícilmente podría yo sentir resentimiento hacia una persona que destrozó la valla de mi jardín con su coche por esquivar a un niño imprudente, y del mismo modo no sería correcto que yo sintiese gratitud hacia mi despistado vecino que la arregló pensando que se trataba de la suya propia. En ninguno de estos dos casos era intención del agente hacerme un bien o un mal, lo uno y lo otro se dieron de forma fortuita.

Las actitudes reactivas llevan aparejado un juicio sobre intenciones porque surgen de ciertas *expectativas* que siempre tenemos a la hora de relacionarnos con los demás. La idea de Strawson es que siempre que nos involucramos en lo que llama “relaciones interpersonales” demandamos buena voluntad por parte de los demás. Generalmente esperamos de los demás que tengan buenas intenciones con respecto a nosotros. Cuando esta expectativa se ve colmada sentimos gratitud, cuando por el contrario no es así sentimos

resentimiento.⁴¹⁶ A participar en este juego social de expectativas, esto es, al estado normal de un agente cuando se relaciona con otros agentes, y que es el estado en que se dan las actitudes reactivas, lo llama Strawson la *actitud participante*. Esta actitud es aquella con la que nos enfrentamos diariamente a los demás en nuestras relaciones interpersonales, en las que, según Strawson, está siempre presente la expectativa de que los demás en principio poseen buenas intenciones para con uno.

Basándose en esta expectativa de buena voluntad que caracteriza a la actitud participante, Strawson consigue distinguir tres grupos de actitudes reactivas en razón de hacia quién se vuelca dicha expectativa. Así, consigue distinguir (a) entre aquellas actitudes reactivas en que esa expectativa se vuelca hacia los demás cuando interactúan conmigo, (b) entre aquellas en que la expectativa se vuelca hacia los demás cuando interactúan entre sí sin formar yo directamente parte de la acción, y por último (c) entre aquellas actitudes en que la expectativa se vuelca hacia mí mismo.

La primera clase de actitudes reactivas la constituyen lo que Strawson llama (a) *actitudes reactivas personales*. Se trata de actitudes tales como el resentimiento o la gratitud, actitudes que reflejan mi expectativa de cierta buena voluntad por parte de los demás en sus acciones con respecto a mí.

La segunda clase de actitudes reactivas la componen lo que Strawson llama (b) *actitudes reactivas morales o análogas vicarias*, pues según Strawson son una transposición exacta de las anteriores al mundo moral. Strawson habla de actitudes tales como la aprobación o la desaprobación moral que reflejan una vez más una expectativa que uno tiene de cierta buena voluntad pero en este caso por parte de los demás para con cualquier otra persona.

⁴¹⁶ Traduzco literalmente “resentment”, pero hay que apuntar que en inglés este término no implica rencor, como sí lo hace en castellano. Pretendo usar “resentimiento” en el mismo sentido en que lo usa Strawson y habrá de leerse así siempre que no se indique lo contrario.

Por último Strawson señala que también poseemos (c) *actitudes auto-reactivas* como el orgullo o el arrepentimiento. En este caso la expectativa de buena voluntad se vuelca hacia uno mismo.

Strawson centra su artículo en las actitudes reactivas personales. Con respecto a las mismas se pregunta cuáles son las condiciones que las hacen razonables o apropiadas, en qué ocasiones está justificado o no que sintamos resentimiento o gratitud hacia alguien. Teniendo en cuenta que dichas actitudes llevan consigo un juicio de intenciones, lo normal es que sintamos resentimiento cuando juzgamos que alguien ha actuado con mala intención y gratitud cuando juzgamos que lo ha hecho con buena intención. Sin embargo esta respuesta no nos dice mucho acerca de las situaciones en que están justificadas nuestras actitudes reactivas. Sí resulta interesante, en cambio, considerar cuáles son los límites a partir de los cuales juzgamos que *no* es apropiado sentirlas. Resulta entonces un conjunto de condiciones que hacen que modifiquemos nuestras actitudes reactivas. Strawson divide este conjunto en dos grupos: condiciones que *excusan* nuestra conducta y condiciones que *eximen* de responsabilidad, según la aclaratoria caracterización que de ellas hace Gary Watson.⁴¹⁷

El primer grupo (el de las excusas) lo integran aquellas ocasiones en que no dejamos de ver al agente como agente plenamente responsable pero sí consideramos que por una determinada acción no ha sido, total o parcialmente, responsable. En estos casos usamos “expresiones tales como ‘No pretendió’, ‘No se dio cuenta’, ‘No sabía’; o también todas aquellas que podrían darnos ocasión de usar la frase ‘No pudo evitarlo’, cuando se apoya en frases tales como ‘Le obligaron’, ‘Fue la única salida’, ‘No le dejaron otra

⁴¹⁷ Véase el apartado “Excusing and Exempting” en Watson 1987, en Fischer & Ravizza 1993: 119-148.

alternativa’, etc.”⁴¹⁸ Normalmente el agente es responsable de sus actos aunque en ese caso concreto, total o parcialmente, no lo fue. En cierto sentido cabe decir que el agente actuó voluntariamente, pero que en realidad no tenía más que una opción realmente válida a la hora de actuar. Estos casos, como señalé en la introducción histórica, ya habían sido apuntados por Aristóteles como ocasiones en las que debemos modificar o suspender nuestros sentimientos morales con respecto al agente. Básicamente se trata de aquellas situaciones en que, siguiendo la terminología aristotélica, el agente actúa por ignorancia o por fuerza mayor.

El segundo grupo (el de las exenciones) consta a su vez de dos subgrupos.

El primer subgrupo (exenciones₁) lo completan las ocasiones en que las condiciones en que actúa el agente no son las normales. En estos casos usamos expresiones “tales como ‘No era él mismo’, ‘Ha estado soportando una gran tensión recientemente’, ‘Estaba actuando bajo sugestión posthipnótica’.”⁴¹⁹ Aquí consideramos que “no deberíamos sentir resentimiento hacia una persona por una acción llevada a cabo por la persona que no es.”⁴²⁰

En cuanto al segundo subgrupo (exenciones₂) lo que falla no son las condiciones en las que el agente lleva a cabo su acción, sino las que constituyen al propio agente. Usamos entonces expresiones del tipo “‘No es más que un niño’, ‘Es un esquizofrénico incurable’, ‘Su mente ha sido metódicamente corrompida’, ‘Eso es un comportamiento puramente compulsivo por su parte’.”⁴²¹ El agente es un enfermo mental o está moralmente subdesarrollado, cuando actúa es él mismo pero es que él mismo no es nunca responsable de sus actos.

⁴¹⁸ Strawson 1962, en Fischer & Ravizza 1993: 50.

⁴¹⁹ Ibidem: 51.

⁴²⁰ Ibidem.

En definitiva, en todo este grupo (exenciones) consideramos que, o bien porque fallan las circunstancias en que se llevó a cabo la acción (exenciones₁), o bien porque el agente no es normal (exenciones₂), el agente no es responsable y suspendemos nuestras actitudes reactivas no sólo en lo que respecta a una acción concreta sino en lo que respecta a cualquier acción que pueda llevar a cabo el agente. En estos casos no puede decirse que el agente actúe voluntariamente.

Según Strawson esta suspensión de las actitudes reactivas que conlleva nuestra actitud participante constituye otro tipo de actitud que él llama *actitud objetiva*. Adoptar la actitud objetiva hacia un agente es verlo como objeto de una “política social”, de tratamiento. La actitud objetiva contiene emociones, pero no actitudes reactivas, no aquellas emociones involucradas en la participación normal con los demás en “relaciones humanas interpersonales”.

II.1.2. La solución de Strawson al problema de la responsabilidad

Llega ahora el momento, una vez desentrañado el juego de nuestras prácticas morales y de la atribución de responsabilidad, de volver al debate sobre el determinismo y la responsabilidad, y preguntarse si el aceptar la verdad de una tesis general del determinismo, debería hacernos repudiar la actitud participante. Esta es la nueva forma que cobra, a la luz del estudio de Strawson sobre las actitudes reactivas, la pregunta sobre la compatibilidad del determinismo y la responsabilidad. De este modo se centra el debate en las consecuencias prácticas del determinismo y se deja de lado la sobreintelectualización característica del tema. Teniendo en cuenta que lo correcto, ante personas que no pueden verse como agentes responsables, es

⁴²¹ Ibidem.

adoptar la actitud objetiva, la pregunta final es pues: ¿Debería el aceptar la verdad de una tesis general del determinismo, hacernos adoptar con respecto a todos los seres humanos la actitud objetiva?

Strawson concede que pensar que esto pueda suceder no es autocontradictorio y así mismo admite que, en ocasiones, somos capaces de adoptar la actitud objetiva con respecto a agentes responsables, frente a personas sanas y maduras “como un refugio de las tensiones de la implicación; o como una ayuda para la política; o simplemente por curiosidad intelectual”.⁴²² Sin embargo Strawson opina que es prácticamente inconcebible un estado de cosas en el que efectivamente adoptásemos de forma continuada para con todos y cada uno de los seres humanos la actitud objetiva. Para defender esta idea Strawson utiliza, en términos de Paul Russell, dos estrategias: una naturalista y otra racionalista.

El argumento naturalista reza que “el compromiso humano de participar en las relaciones interpersonales comunes está profundamente arraigado y es demasiado incondicional para tomar en serio la idea de que una convicción teórica general pueda cambiar así nuestro mundo”⁴²³ o incluso que “una objetividad de actitud interpersonal sostenida, y el aislamiento humano que entrañaría, no parece algo de lo que los seres humanos puedan ser capaces, ni aun cuando una verdad general fuese una base teórica para ello”.⁴²⁴ En definitiva, según Strawson es imposible que adoptemos universalmente la actitud objetiva porque *no está en nuestra naturaleza* poder hacerlo, y aunque lo estuviese sería tan contrario a nuestra forma normal de comportarnos, que una exigencia de orden teórico no sería nunca capaz de modificar nuestras prácticas morales habituales, y hacernos suspender nuestras actitudes reactivas.

⁴²² Ibidem: 52. En el apartado *III.1.2. (La actitud objetiva y la investigación teleológica)* retomaré este punto que parece contradecir muchas de las ideas que Strawson despliega en su artículo.

⁴²³ Ibidem.

⁴²⁴ Ibidem.

En cuanto a la estrategia racionalista Strawson señala que la actitud objetiva, salvo excepciones en que la adoptamos para con alguien “normal”⁴²⁵, es “consecuencia de que veamos al agente como *incapacitado* en alguno o en todos los aspectos en lo que respecta a las relaciones interpersonales ordinarias”.⁴²⁶ Strawson piensa que cuando adoptamos la actitud objetiva no es “por una convicción teórica que expresaríamos como ‘determinismo en este caso’, sino como consecuencia de que abandonamos, por diferentes razones en diferentes casos [excusas y exenciones], las actitudes interpersonales ordinarias”.⁴²⁷ Principalmente esta estrategia racionalista es la que seguiré más adelante para reforzar la posición compatibilista.

II.1.3. La racionalidad de la responsabilidad moral

Pero Strawson no considera que la pregunta más importante a la que hay que responder, respecto al problema de la responsabilidad moral, sea la que se refiere a si es posible o no un estado de cosas en el que sostendríamos de forma continuada una actitud objetiva para con todos y cada uno de los seres humanos, sino “una pregunta acerca de qué sería *racional* hacer si el determinismo fuera cierto, una pregunta acerca de la justificación racional de nuestras actitudes interpersonales ordinarias en general”.⁴²⁸

Este cambio de perspectiva en la discusión es importante porque la respuesta a esta pregunta exige la especificación de algún tipo de criterio de racionalidad. Aunque fuera posible, en razón de la verdad de una tesis general del determinismo, adoptar para toda persona la actitud objetiva, resulta que tal vez no sería racional hacerlo. Ante este nuevo dilema Strawson una vez más da

⁴²⁵ Véase la nota 422 y el texto al que corresponde.

⁴²⁶ Strawson 1962, en Fischer & Ravizza 1993: 55.

⁴²⁷ Ibidem.

⁴²⁸ Ibidem.

dos respuestas: una siguiendo una estrategia naturalista y otra siguiendo una estrategia racionalista.

La primera de las respuestas es idéntica a aquella que da Strawson con respecto al problema de si sería posible o no universalizar la actitud objetiva, a saber: tenemos un compromiso profundo con la actitud participante, “ese compromiso es parte del entramado general de la vida humana, no algo que pueda ser revisado en casos particulares sin revisarse con ello el entramado general”.⁴²⁹ Así, no está en nuestra naturaleza el poder de adoptar universalmente la actitud objetiva y “es *inútil* preguntar si no sería racional para nosotros hacer lo que no está en nuestra naturaleza (ser capaces de) hacer”.⁴³⁰

La segunda de las respuestas, en cambio, introduce un criterio de racionalidad no presente con anterioridad en la discusión acerca del concepto de responsabilidad moral y que a mi juicio es muy interesante: “podríamos escoger racionalmente sólo a la luz de una valoración de las ganancias y pérdidas para la vida humana, su enriquecimiento o empobrecimiento; y la verdad o la falsedad de una tesis general del determinismo no podría menoscabar la racionalidad de *esta* elección”.⁴³¹

Aquí, una vez más, Strawson está convirtiendo el debate metafísico en debate moral, no así en sus argumentos naturalistas que, como intentaré mostrar más adelante, también acaban constituyendo una sobreintelectualización de los hechos. Salvo en contadas ocasiones⁴³² este criterio pragmatista de racionalidad ha sido soslayado de la discusión acerca del concepto de responsabilidad moral, aquí defenderé que, por el contrario, es el más importante. También discutiré en breve hasta qué punto es coherente por parte de Strawson compaginar una estrategia naturalista y otra racionalista.

⁴²⁹ Ibidem: 55.

⁴³⁰ Ibidem: 60.

⁴³¹ Ibidem: 55-56.

⁴³² En concreto Ayer y Bennett destacan ambos la importancia de este argumento en Ayer 1980 y Bennett 1980, publicados en Van Stratten 1980.

De momento voy a conformarme, a modo de síntesis para este apartado, con dejar bien claro cuáles son los diferentes tipos de argumento que presenta Strawson en su artículo en contra de las consecuencias prácticas del incompatibilismo. De aquí en adelante hablaré de tres argumentos de Strawson contra la idea de que, dado el determinismo, no tiene sentido el concepto de responsabilidad moral:

(1) El *naturalista*, según el cual nunca sería posible que adoptásemos universalmente la actitud objetiva aun cuando el determinismo resultase ser cierto y por lo tanto no tiene sentido preguntarse si sería o no racional adoptar esa actitud de ser cierto el determinismo.

(2) El *normativista*, según el cual adoptar la actitud objetiva lleva consigo siempre, por las razones estudiadas (excusas y exenciones) y no siendo en ningún caso nuestra convicción teórica de que el determinismo es cierto una de esas razones, el reconocer al otro como alguien inadecuado para provocar en nosotros las actitudes reactivas presentes normalmente en nuestras relaciones interpersonales.

(3) El *pragmatista*, según el cual adoptar universalmente la actitud objetiva supondría una pérdida importante para nuestras vidas y por lo tanto no sería racional hacerlo.

Nótese que el argumento naturalista sirve al mismo tiempo para responder a dos preguntas: una sobre la posibilidad de adoptar universalmente la actitud objetiva y otra sobre la racionalidad de este proceder. Esto es así porque la respuesta naturalista a la primera pregunta hace irrelevante la

segunda. Por el contrario dentro de una estrategia racionalista es necesario responder a ambas. Cuando hable, pues, de la estrategia naturalista de Strawson no será necesario especificar más de un argumento. Por el contrario, cada vez que nombre la estrategia racionalista habrá que tener presente que ésta lleva consigo dos argumentos: el que he llamado normativista y el que he llamado pragmatista. Es importante no confundir estrategias y argumentos.

II.1.4. En torno al utilitarismo

Hasta aquí sólo he tratado los argumentos de Strawson contra el incompatibilismo, debo retomar ahora la crítica al utilitarismo. Ésta tendría dos aspectos: uno metafísico y otro moral. Hasta este momento sólo he rechazado el argumento utilitarista en el plano metafísico por juzgar que en absoluto responde a los problemas planteados por el incompatibilista. Ahora conviene analizar el aspecto moral del argumento utilitarista. Una vez más es ese aspecto, el moral, el que interesa a Strawson, pues éste considera que aquello por lo que es insostenible la postura utilitarista es porque postula una objetividad de actitud universal que excluye elementos esenciales del concepto de responsabilidad moral.

Según Strawson, las actitudes reactivas que nos llevan a adoptar nuestras prácticas sociales comunes de premio y castigo son el resultado de ver al otro como miembro de *una comunidad moral* y “hablar en términos de utilidad social sin más es dejar fuera algo vital en nuestra concepción de dichas prácticas”.⁴³³ El utilitarista olvida que nuestras prácticas sociales de premio y castigo, y las reacciones que llevan consigo “realmente *son* expresiones de nuestras aptitudes morales y no meros patrones que calculadoramente

⁴³³ Strawson 1962, en Fischer & Ravizza 1993: 63.

empleamos con propósitos regulativos”.⁴³⁴ Esa es la laguna del utilitarista, del compatibilista clásico, que llenan las actitudes reactivas, pues éstas son esenciales “para el sentido de una comunidad moral”.⁴³⁵

En “Freedom, Blame and Moral Community”, Lawrence Stern explica que el utilitarista no está equivocado por pensar que el intentar reformar a un sujeto es una de las cosas que justifica nuestras prácticas de premio y castigo, sino por entender esta relación entre la reforma de una persona y su castigo de manera casi conductista. No es equivalente lo que hacemos con las personas para mejorarlas moralmente, a lo que hacemos con los animales en vistas a cambiar su comportamiento. Cuando estamos en presencia de seres humanos “estamos apelando a principios con los cuales la otra persona presuntamente puede funcionar con éxito por sí misma; también, esperamos de ella que vea la aplicación de los principios relevantes en cada caso. Actuamos con la esperanza de una comunidad racional y moral: nos comprometemos en un diálogo.”⁴³⁶ Entender la presencia de la desaprobación moral en términos utilitaristas es dar por hecho que no hay ningún diálogo posible con la persona condenada moralmente, sino tan sólo la posibilidad de sacar partido de ella socialmente mediante su reforma, como cuando entrenamos a un perro para que detecte droga en las maletas en los aeropuertos o para acostumbrarlo a que miccione siempre en la calle y no en nuestra moqueta. Insistiendo en esta idea Susan Wolf recuerda que “normalmente no elogiamos o culpamos a las demás personas porque, a resultas de una deliberación práctica, hayamos llegado a la conclusión de que sería interesante hacerlo. Más bien, alabamos o culpamos a las personas como una expresión natural de respuestas naturales a lo que vemos hacer a la gente”.⁴³⁷

⁴³⁴ Ibidem: 66.

⁴³⁵ Stern, L. 1994: 79

⁴³⁶ Ibidem.

⁴³⁷ Wolf 1981, en Fischer & Ravizza 1993: 104.

El utilitarista escoge como criterio de racionalidad la utilidad práctica de ciertos comportamientos sociales que conlleva la existencia de la responsabilidad moral, pero el que tal elección sea posible depende de que demos por hecho algo que es precisamente lo que está en juego, a saber: los seres humanos podemos escoger a voluntad si queremos o no adoptar nuestras diferentes actitudes reactivas, o lo que es lo mismo, vivimos permanentemente en una actitud objetiva (lo cual es bien discutible) que nos permite, una vez hemos calibrado las ganancias que conllevará desde el punto de vista de la utilidad social, escoger o no el comportarnos como seres responsables, o juzgar o no a los demás como tales seres responsables. Pero esto último parece claramente incompatible con la idea de que los seres humanos vivimos en una comunidad moral tal y como ésta ha sido esbozada más arriba: un lugar en que funcionamos por medio de razones y principios en que basamos en gran medida nuestros sentimientos.⁴³⁸

Así pues, como mínimo cabe exigirle al utilitarista que refine su argumento, pues en él aparecen ciertos presupuestos que necesitan un estudio más atento. Más allá de ese mínimo, Strawson como naturalista tiene muy claro que el argumento utilitarista es manifiestamente falso, pues falso es aquello que da por hecho, esto es: que podemos escoger a voluntad sentir o no las actitudes reactivas como si estuviésemos permanentemente inmersos en la actitud objetiva. Probablemente no es del todo falso afirmar que precisamente Strawson insiste en la estrategia naturalista para no dar ninguna posibilidad a la vía utilitarista. Aún así parece que es suficiente, para deslegitimar la postura utilitarista, mostrar que obvia aspectos demasiado importantes del debate acerca de la responsabilidad moral y para esto, como se acaba de ver, basta con una estrategia racionalista como la de Lawrence Stern.

⁴³⁸ En el último capítulo del trabajo (III-4.) desarrollaré con más precisión este punto.

La mejor prueba de esto último la da Robert Jay Wallace⁴³⁹ que recoge en parte los argumentos de Strawson aunque los presenta de forma más ordenada y desde una perspectiva más cognitivista. Según Wallace hay tres razones por las cuales puede considerarse que la estrategia utilitarista es incorrecta.

La primera de ellas recoge la idea de Strawson, Stern y Wolf de que el utilitarismo deja fuera aspectos esenciales de lo que se suele entender por responsabilizar moralmente, pues sólo tiene en cuenta el castigo como comportamiento tipo, como patrón de la atribución de responsabilidades, y es esto mismo lo que lleva al utilitarismo a caer en descripciones excesivamente conductistas olvidando que “considerar a las personas moralmente responsables de sus acciones es verlas como abiertas a algún tipo de valoración profunda (...) que es presumiblemente moral en su calidad y fuerza.”⁴⁴⁰ En efecto, cuando hacemos una atribución de responsabilidad no sólo estamos sancionando o persiguiendo una futura sanción, sino que estamos sintiendo algo, y eso es lo que revela Strawson: atribuir responsabilidad no es sólo algo que hacemos, sino también algo que, en cierto sentido, padecemos. He ahí por qué es indispensable hablar de actitudes: porque la responsabilidad moral no incluye exclusivamente preceptos, sino también emociones, y esto el utilitarista (tal y como lo caracteriza Strawson) lo olvida por completo.

Este olvido lleva al utilitarista a enfrentarse con un segundo problema, a saber, el de que el utilitarismo sólo permite adscripciones de responsabilidad con carácter disuasorio, pero no con el carácter de enjuiciamiento de una acción concreta que suelen tener nuestras atribuciones de responsabilidad normalmente. Así lo explica Susan Wolf: “(...) aunque la aprobación y la desaprobación moral *puedan* ser consideradas como tipos de acciones, nuestra

⁴³⁹ Wallace 1998. Véase a este respecto el apartado “3.1 Responsibility, Blame and Moral Sanction”.

⁴⁴⁰ Wallace 1998: 53-54.

experiencia ordinaria de estos fenómenos nos anima a considerarlos como expresiones de cierto tipo de juicio. [...] Justificar la alabanza y la culpa en la forma sugerida por el optimista [léase utilitarista] es dejar sin explicar tales juicios acerca del mérito individual.”⁴⁴¹ En efecto, la perspectiva utilitarista convierte nuestra desaprobación moral en algo que no busca sino provocar cierto efecto, a saber: que no vuelva a repetirse aquello que juzgamos reproable.⁴⁴² El utilitarista, al reducir la responsabilidad moral a una herramienta para el progreso social, no tiene en cuenta lo que nuestra atribución de responsabilidad tiene de juicio hacia una persona concreta (que puede ser uno mismo), de tal forma que dicho juicio, desde la perspectiva utilitarista “no se dirige exclusivamente hacia el agente individual que ha hecho algo moralmente incorrecto, sino que responsabiliza a cualquiera susceptible de ser influenciado por nuestras respuestas”.⁴⁴³ La idea de la indignación y el castigo como medios disuasorios deja fuera nuestra intuición de que responsabilizar moralmente implica que hay alguien concreto que ha cometido un acto reproable y que debería reparar, si posible, su falta. En efecto, si nos ponemos en una perspectiva utilitarista radical podemos pensar que es indiferente que castigemos a un inocente o a un culpable a efectos de la disuasión de futuros delincuentes, bastaría con hacerle creer a la sociedad que castigamos a un culpable para que el castigo cobrase su carácter disuasorio (es irrelevante la calidad moral real de un chivo expiatorio). La definición utilitarista de la responsabilidad moral no deja claro el límite entre quienes son responsables y quienes no lo son, y sólo una cierta “preocupación benéfica”⁴⁴⁴ podría ser un argumento para utilizar como ejemplo disuasor más bien a un culpable que a un inocente. Cuando atribuimos responsabilidad llevamos con

⁴⁴¹ Wolf 1981, en Fischer & Ravizza 1993: 104.

⁴⁴² Más adelante se verá que ciertamente éste es también un aspecto a tener en cuenta en el juicio moral, pero no el principal.

⁴⁴³ Wallace 1998: 57.

⁴⁴⁴ Ibidem.

nosotros algo diferente de un propósito educador: “aunque una preocupación benéfica nos llevase a sancionar únicamente a personas efectivamente culpables, la actitud expresada por tales sanciones no sería la actitud expresada por nuestras respuestas morales reales a las malas acciones.”⁴⁴⁵ Cuando responsabilizamos moralmente nos estamos ocupando de una persona concreta, no de un “mal ejemplo” abstracto: “La evaluación moral es importante porque incide en el comportamiento *vis à vis* de las personas evaluadas. Debemos encarcelar, amenazar o culpar a los delincuentes, no a su comportamiento; y nuestra conducta es racional sólo en tanto en cuanto sea compatible con lo que hemos descubierto acerca de los delincuentes a través de su comportamiento.”⁴⁴⁶ En fin, estamos una vez más en el problema anterior porque estas consecuencias indeseables del utilitarismo son fruto de una mala definición y una mala comprensión del concepto de responsabilidad moral, tal y como se vio más arriba.

Por fin, en tercer lugar Wallace retoma la idea en la que vengo insistiendo desde el principio de este capítulo, según la cual el utilitarismo no deja fuera en absoluto al incompatibilismo: “Si, como mantiene, la atribución de responsabilidades es una disposición para reprender y sancionar en vías a influenciar el comportamiento de las personas, es muy difícil ver por qué podría haber una profunda o persistente tentación de pensar en la responsabilidad moral en términos incompatibilistas.”⁴⁴⁷

En definitiva, desde una argumentación strawsoniana no se puede por menos que señalar que el utilitarismo tiene poco que decir dentro del debate acerca del concepto de responsabilidad moral. Como dice Wallace, el utilitarismo “hace el compatibilismo *tan* plausible que no puede ser correcto como explicación de lo que realmente hacemos cuando consideramos

⁴⁴⁵ Ibidem.

⁴⁴⁶ Cummins 1980: 212.

⁴⁴⁷ Wallace 1998: 58.

moralmente responsable a una persona”.⁴⁴⁸ Por ello, y en razón de los argumentos que han sido desarrollados en este apartado, consideraré que, si bien será posible rescatar en algún momento alguna intuición utilitarista, no será necesario ya atender a la perspectiva utilitarista como una vía interesante a la hora de explicar en qué consiste la responsabilidad moral y desde luego consideraré que en ningún caso es el punto de vista utilitarista una herramienta válida para defender el compatibilismo.

II-2. La responsabilidad: ¿Nace o se hace?

«¿Quién me ha hecho a mí? (...) ¿Pues de dónde me vino a mí el querer mal y no querer el bien? (...) ¿Quién puso esta voluntad dentro de mí? (...) Y si la puso el diablo, ¿quién hizo al diablo? Y si el ángel bueno se hizo demonio por su mala voluntad, ¿cómo llegó a poseer esa mala voluntad que le hizo demonio (...)?»⁴⁴⁹

Como se ha visto más arriba, Strawson da cuatro argumentos para combatir el incompatibilismo. Dos de estos argumentos están destinados a mostrar que, en cierto sentido, no es posible adoptar una posición incompatibilista. Los otros dos argumentos pretenden mostrar que, aún de ser posible adoptar una posición incompatibilista, no sería racional hacerlo. De estos cuatro argumentos dos de ellos se dan dentro de una estrategia argumentativa de corte naturalista y los otros dos dentro de una estrategia argumentativa de corte racionalista.

Dentro de la estrategia naturalista los argumentos serían los siguientes:

- Aún de ser cierto el determinismo, no podemos mantener con sentido una posición incompatibilista, pues dicha

⁴⁴⁸ Ibidem.

⁴⁴⁹ Agustín 1990: 173.

postura implica una permanente objetividad de actitud que por naturaleza es imposible.

- Aún de ser cierto el determinismo, no tiene sentido preguntarse si sería racional o no asumir la objetividad de actitud que postula el incompatibilista, pues no está en nuestra naturaleza la capacidad para hacer tal asunción.

Dentro de la estrategia racionalista los argumentos serían los siguientes:

- Aún de ser cierto el determinismo, la posición incompatibilista, en tanto postula una permanente objetividad de actitud, no tiene sentido en el marco de nuestras relaciones interpersonales normales, pues la actitud objetiva sólo tiene sentido como excepción a nuestra actitud normal en las relaciones interpersonales, a saber, la actitud participante.

- Aún de ser cierto el determinismo, no sería racional adoptar de forma permanente la actitud objetiva que postula el incompatibilismo, pues nuestra vida es mejor dentro de la actitud participante.

Claramente se ve que en la estrategia naturalista los argumentos tienen un aire esencialista. Más arriba señalé que en realidad sólo hay un argumento naturalista, pues la respuesta a la pregunta sobre la racionalidad del incompatibilismo desde la estrategia naturalista, tal y como la desarrolla Strawson, deja de ser necesaria. En cuanto a la estrategia racionalista tiene un aire pragmatista, especialmente en lo que se refiere al segundo argumento, pues lleva consigo una idea de aquello que sería racional que depende de lo que consideremos mejor para nuestras vidas.

¿Hasta qué punto tiene sentido que Strawson emplee al mismo tiempo argumentos esencialistas y argumentos pragmatistas? La respuesta a esta pregunta es lo que me va a ocupar en este apartado. Así, primero estudiaré los problemas de cada una de estas dos estrategias por separado, y veré hasta qué punto es realmente necesario que Strawson desarrolle ambas. A continuación estudiaré si esa iniciativa de Strawson es un acierto y debería ser conservada, o por el contrario es un problema de su teoría al que se debería tratar de hacer frente.

II.2.1. Problemas con la estrategia naturalista

En su indispensable artículo “Strawson’s Way of Naturalizing Responsibility”, Paul Russell comienza señalando lo siguiente: “Strawson no logra distinguir dos formas o modos muy diferentes de naturalismo y se ve obligado por la naturaleza de sus propios objetivos (por ejemplo la refutación del Pesimista [léase incompatibilista]) a abrazar la forma de naturalismo más fuerte y de lejos menos plausible. Sobre esta base concluyo que (...) concretamente a través de las líneas que él sugiere no podemos naturalizar la responsabilidad.”⁴⁵⁰

Paul Russell señala que hay dos formas de entender al incompatibilista: o bien como un incompatibilista tipo o bien como un incompatibilista ejemplar. El *incompatibilista tipo* sostendría que, de ser cierto el determinismo, no estaría justificado que estemos expuestos a las actitudes reactivas y por ello deberíamos deshacernos de esa clase de actitudes. El *incompatibilista ejemplar* se limitaría a señalar que, si encontrásemos que nada justifica las actitudes reactivas dadas nuestras circunstancias, entonces seríamos capaces de dejarlas de lado. Así pues, la diferencia entre ambos tipos de incompatibilismo estriba

⁴⁵⁰ Russell, P. 1992: 288.

en que el primero de ellos sostiene que debemos dejar las actitudes reactivas, y el otro considera que si llegásemos a ser conscientes de la ausencia de justificación de las actitudes reactivas entonces seríamos capaces de dejarlas.

Pues bien, del mismo modo que hay dos formas de entender el incompatibilismo, hay dos formas de interpretar el naturalismo de Strawson. La primera de ellas, la de un *naturalismo tipo*, rezaría que “regular e inevitablemente encontramos circunstancias bajo las cuales (...) [el comportamiento responsable] es plenamente apropiado y razonable y, por lo tanto, los sentimientos y experiencias (...) [de la responsabilidad] continuarán siendo una parte de la vida humana de la que no podemos escapar”.⁴⁵¹ La segunda consistiría en una suerte de *naturalismo ejemplar* según el cual ningún tipo de razón podría hacernos jamás abandonar nuestro comportamiento como agentes responsables ni las actitudes reactivas que conlleva.

Lo primero que se advierte es que, de las dos formas de naturalismo, la primera de ellas, por sí sola, es incapaz de rebatir las dos formas de incompatibilismo. En efecto, la idea de que debemos mantener nuestras actitudes reactivas porque están justificadas por nuestras circunstancias (*naturalismo tipo*), no contradice en absoluto la idea de que de no estarlo podrían y deberían ser dejadas de lado (*incompatibilismo ejemplar*). Así pues uno se ve forzado a abrazar el segundo tipo de naturalismo, el más fuerte, que afirma de forma esencialista que nada será nunca capaz de hacernos rechazar nuestras actitudes reactivas porque estamos indisolublemente unidos a ellas independientemente de lo que diga la razón, de lo que creamos o no con respecto al sentido de nuestras actitudes reactivas.

De este modo sólo el naturalismo ejemplar aparece como una herramienta útil a la hora de enfrentarse con el incompatibilismo, sin embargo esta herramienta tampoco está exenta de problemas: “(...) está lejos de ser

⁴⁵¹ Ibidem: 294.

obvio –es más, parece sencillamente falso- que seamos constitucionalmente incapaces de dejar de abrazar o sentir (...) [las actitudes reactivas] en circunstancias en las que creemos que no resultan apropiadas o que no vienen a cuento.”⁴⁵² Y además, aunque admitiéramos que esto no es así, tal vez el incompatibilista tipo se vería obligado a aceptar que no hay forma de deshacerse de nuestras actitudes reactivas, pero seguiría pensando que el naturalista debería demostrarle que hay alguna razón que justifique el que esto sea así. Para el incompatibilista tipo no basta con afirmar que nada podría hacernos dejar de lado la responsabilidad, sino que también es necesario mostrar que es racional no hacerlo (dicho de otro modo: para el incompatibilista tipo, independientemente de que podamos o no obrar en consecuencia, sí es relevante preguntarse por la racionalidad de la responsabilidad).

Así pues, la estrategia naturalista de Strawson parece insuficiente para rebatir los argumentos del incompatibilista, y no obstante da la impresión de que tanto el incompatibilismo tipo como el incompatibilismo ejemplar están equivocados, pues el primero “tergiversa la manera en que nuestra disposición a (...) [la responsabilidad] está enclavada en nuestra naturaleza humana”⁴⁵³ y el segundo se equivoca al sostener “que las circunstancias nunca son tales que (...) [nuestra responsabilidad] está en orden o justificada”.⁴⁵⁴ El problema es que, como se ha visto, sólo el naturalismo ejemplar da cuenta del primero de estos problemas y sólo el naturalismo tipo podría hacerlo vagamente del segundo.

Mi impresión, siguiendo el razonamiento de Paul Russell, es que Strawson aspira sobre todo a deshacer una forma incorrecta de entender el concepto de responsabilidad moral que, como ya se dijo, incurre en una

⁴⁵² Ibidem: 295.

⁴⁵³ Ibidem.

⁴⁵⁴ Ibidem.

sobreintelectualización de los hechos. Por lo tanto parece que Strawson intenta rebatir principalmente el incompatibilismo tipo, y por ello propone una forma de naturalismo muy fuerte, el ejemplar, que es la que más directamente se opone al incompatibilismo tipo. No obstante, consciente de que es discutible, como se dijo hace un momento, que bajo ninguna circunstancia seamos capaces de abandonar nuestras actitudes reactivas, Strawson se ve obligado a refinar su naturalismo ejemplar y a sumarle una serie de consideraciones y argumentos de carácter racionalista. Así pues, da la impresión de que Strawson recurre a la estrategia racionalista para atar los cabos que su argumento naturalista deja sueltos al intentar falsar el incompatibilismo. Por ello hay que concluir que el éxito de la argumentación de Strawson reposa en gran medida en la posibilidad de que una estrategia naturalista y una estrategia racionalista sean compatibles. Así, es indispensable ver si no hay algún problema que haga que ambas estrategias sean incompatibles, pues de serlo la teoría de Strawson quedaría coja. Pero antes de ocuparme de esto, igual que acabo de estudiar algunos problemas de la estrategia naturalista, debería ver si no hay alguno del lado de la estrategia racionalista.

II.2.2. Problemas con la estrategia racionalista

El problema que Paul Russell detecta en la estrategia racionalista de Strawson es que en el marco de la misma usa indistintamente dos términos que no son equiparables, a saber: “anormalidad” e “incapacidad”. En efecto, Strawson nos dice que la actitud objetiva viene dada porque consideramos al otro como excluido de nuestras relaciones interpersonales ordinarias “por una anormalidad psicológica profundamente arraigada —o simplemente por ser un niño. Pero no puede ser consecuencia de ninguna tesis que no sea ella misma

contradictoria que la anormalidad es la condición universal.”⁴⁵⁵ Bien, Paul Russell señala con acierto que lo que en realidad impulsa la actitud objetiva no es el que un determinado agente sea considerado anormal (sin ir más lejos el niño, según Strawson, está excluido de las relaciones interpersonales ordinarias por estar “moralmente subdesarrollado”, no por ser anormal), sino que más bien nuestra actitud objetiva es, en palabras de Strawson, “consecuencia de que veamos al agente como *incapacitado* en alguno o en todos los aspectos en lo que respecta a las relaciones interpersonales ordinarias”.⁴⁵⁶ Según Paul Russell las dos afirmaciones de Strawson que acabo de citar no son equiparables, y tampoco lo son sus efectos sobre las ideas del incompatibilista, pues éste no pretende decir en ningún momento que si el determinismo es cierto entonces todos somos anormales sino, más bien, que si lo es, entonces estamos todos moralmente incapacitados. Si esto es así el argumento normativista de Strawson pierde toda su fuerza, pues es bien cierto que es autocontradictorio pretender que la anormalidad es condición universal, pero ya no lo es defender que estamos universalmente incapacitados moralmente, esto es, que no somos objetos apropiados para las actitudes reactivas sino tan sólo para la actitud objetiva. El ejemplo que suministra Paul Russell para mostrar hasta qué punto no pueden solaparse incapacidad y anormalidad como si de una misma cosa se tratase, es el siguiente: “Imaginemos, por ejemplo, que se esparce una terrible enfermedad o una mutación genética que afecta al cerebro y con ello destruye nuestras capacidades relevantes [para nuestras relaciones interpersonales ordinarias]. Claramente, en esa situación no hay correspondencia o equivalencia extensional entre el «anormal» y el «incapacitado». Al contrario, la persona «normal» sería incapacitada y la persona «anormal» tendría las capacidades requeridas [para nuestras relaciones interpersonales ordinarias].

⁴⁵⁵ Strawson 1962, en Fischer & Ravizza 1993: 53-54.

⁴⁵⁶ Ibidem: 55.

Dado esto, nuestras actitudes reactivas serían inapropiadas en el caso normal y apropiadas en el caso anormal.”⁴⁵⁷

Este ejemplo muestra que la estrategia racionalista, tal y como la desarrolla Strawson, depende de la naturalista tanto como ésta lo hace de la primera, pues para que el argumento normativista funcione es necesaria una idea de incapacidad que sea equivalente a la de anormalidad, y esta idea es la que aporta el argumento naturalista al mostrarnos que, por naturaleza, no es posible que perdamos nuestra capacidad para funcionar haciendo atribuciones de responsabilidad, lo cual equivale a decir que no sólo no es normal sino ni tan siquiera posible que suspendamos nuestra capacidad para hacer dichas atribuciones.

II.2.3. ¿Incompatibilidad de las dos estrategias?

En los apartados anteriores se ha terminado llegando a una conclusión en principio inesperada, a saber: que la estrategia naturalista y la racionalista, tal y como las desarrolla Strawson, son interdependientes. Digo que esta conclusión resulta en principio inesperada porque, como dije más arriba, el naturalismo strawsoniano responde a un tipo de naturalismo esencialista, lo cual parece claramente ir en contra de una estrategia normativista. Así lo ve también Paul Russell: “Lo que es particularmente molesto en la estrategia naturalista de Strawson, expresado en términos generales, es que arroja dudas sobre nuestra habilidad o capacidad para contener o controlar nuestra vida emocional de acuerdo con los dictados de la razón.”⁴⁵⁸

Ahora, tal vez me he precipitado a la hora de sacar conclusiones y el que la estrategia racionalista y la estrategia naturalista sean interdependientes en el artículo de Strawson debería más bien ser tomado como un síntoma de

⁴⁵⁷ Russell, P. 1992: 299.

que éste no logra desarrollar correctamente ninguna de las dos estrategias, y no como la demostración de que efectivamente no puede combatirse el incompatibilismo sin ambas. Así pues, no estaría de más analizar a continuación hasta qué punto es cierta la intuición de que ambas estrategias son incompatibles, pues de serlo no se podría por menos que concluir que hay que optar por una de ellas para argumentar a favor de la articulación de la responsabilidad moral que defiende Strawson, y que será lo que intentaré desarrollar de forma pormenorizada en la tercera parte de este trabajo. Se debería pues, en ese caso, admitir que la forma en que argumenta Strawson en contra del incompatibilismo no es plenamente satisfactoria, y corregirla hasta lograr una argumentación sólida en que asentar la caracterización del hecho de la responsabilidad moral que hace Strawson, en torno a la actitudes reactivas. Y bien, dicho esto pasemos a ver hasta qué punto la estrategia naturalista y la racionalista son compatibles.

Como se ha visto, un naturalismo fuerte como el que propone Strawson acaba por dejar de lado un hecho que él mismo describe muy bien en su artículo, a saber: existen ocasiones en que somos capaces de suspender nuestras emociones reactivas porque consideramos que hay algo que excusa o exime de responsabilidad a un determinado agente. Para hacer encajar este hecho con su argumento naturalista Strawson formula otro argumento, el normativista, según el cual podría considerarse que ese momento en que suspendemos nuestras emociones reactivas es también natural en el sentido de que ese abandono de las mismas que implica adoptar la actitud objetiva es la reacción natural a nuestra percepción de una cierta anormalidad en la responsabilidad del agente. Se ve que, en cualquier caso, el que esta unión de los dos argumentos funcione sólo puede ser posible por dos vías: o bien

⁴⁵⁸ Ibidem: 297.

naturalizando el argumento normativista, o bien racionalizando el argumento naturalista.

Si naturalizamos el argumento normativista, “incapacidad” y “anormalidad” se hacen equiparables lo cual lleva a considerar que, del mismo modo que es natural nuestra capacidad para actuar responsablemente, es natural, en algunos seres humanos, la incapacidad para actuar responsablemente y, lo que es más, es también natural nuestra disposición a adoptar la actitud objetiva cuando damos con alguno de estos seres humanos que por naturaleza son incapaces de actuar responsablemente. Las consecuencias de adoptar esta perspectiva, que además contradice claramente nuestra experiencia de la responsabilidad, serían sin duda funestas. Así pues no sería posible naturalizar de este modo el argumento normativista, y sin embargo esta parece ser la única forma de hacerlo desde la teoría de Strawson.

Si decidiésemos racionalizar el argumento naturalista no podríamos en ningún caso sostener la tajante afirmación de Strawson de que por naturaleza es imposible mantener de forma permanente la actitud objetiva pues, desde la perspectiva normativista, adoptar o no la actitud objetiva depende de encontrar o no razones para adoptarla, y tal vez algún día podríamos llegar a considerar que el determinismo es una buena razón para adoptarla de forma permanente (una razón que eximiría de responsabilidad).⁴⁵⁹

Así pues, ninguna de las formas en que podríamos conciliar el argumento normativista y el naturalista es satisfactoria ya que: o bien caemos en una definición de la racionalidad que se haría en términos casi conductistas, o bien en una forma de naturalismo tan desnaturalizada que dejaría de tener sentido decir de la responsabilidad que es algo natural. No se puede, pues, por menos que concluir que no es posible llevar adelante ambas estrategias

⁴⁵⁹ Un poco más adelante me ocuparé de nuevo de estos argumentos y los analizaré con más cuidado, de momento tan sólo son necesarios para percibir la incompatibilidad de la estrategia racionalista y la naturalista en la forma en que Strawson las desarrolla.

simultáneamente en la forma en que lo hace Strawson. Ahora, ¿quiere esto decir que ambas estrategias son incompatibles? Creo que no.

Como dije más arriba, el argumento naturalista de Strawson más débil (el del naturalismo tipo) tiene el problema de que nos informa acerca de cómo están las cosas, pero no nos dice nada acerca de cómo podrían estar. Sin embargo Strawson aspira a decir no sólo que las cosas están de tal manera que es normal que actuemos según las normas que entraña la existencia de responsabilidad moral, sino que las cosas no podrían, ni podrán estar nunca, de otra manera que como efectivamente están. Por otra parte, la estrategia racionalista, si equipara “incapacidad” y “anormalidad”, no acaba de explicar correctamente que las cosas estén como están, tan sólo dice que no deberían cambiar. Así pues, creo que lo mejor sería reunir ambas estrategias limitando el alcance de cada una de ellas a aquello en lo que es útil, a saber: la estrategia naturalista para enseñarnos cual es el fundamento de la responsabilidad, y la estrategia racionalista para mostrarnos por qué, dada la responsabilidad, es más racional dejar las cosas como están.

Además, intuyo que el argumento pragmatista, al que temporalmente he dado de lado, podría ser de gran ayuda como trasfondo para ambas estrategias. No puedo avanzar más de momento, voy pues, para empezar, a investigar cual es el fundamento de nuestra “natural” percepción de los seres humanos como seres responsables.

II-3. El yo y su autoconciencia: fenomenología del sujeto responsable

“En aqueste, pues, del sol / ya frenesí, o ya delirio, / nació
Segismundo, dando / de su condición indicios, / pues dio la muerte a su
madre, / con cuya fiera dijo: / Hombre soy, pues que ya empiezo / a
pagar mal beneficios. / Yo, acudiendo a mis estudios, / en ellos y todo
miro / que Segismundo sería / el hombre más atrevido, el príncipe más
cruel / y el monarca más impío, / por quien su reino vendría / a ser

parcial y diviso, / escuela de las traiciones / y academia de los vicios
(...).”⁴⁶⁰

Gran parte de las críticas que le han sido dirigidas a Strawson, incluso por parte de sus simpatizantes, han surgido porque no consigue dar un argumento convincente que justifique su tajante afirmación de que “es *inútil* preguntar si no sería racional para nosotros hacer lo que no está en nuestra naturaleza (ser capaces de) hacer”.⁴⁶¹ En efecto, según Ayer el único argumento que esgrime Strawson para justificar su idea de que nada podría hacernos abandonar nuestras actitudes reactivas es que “siempre que adoptamos la actitud objetiva hacia algún agente, tanto en lo que respecta a algunas como en lo que respecta a todas sus acciones, es siempre por una razón especial”.⁴⁶²

II.3.1. La posibilidad de lo imposible

La idea de Ayer es que es un error considerar que es irrelevante contar con una noción clara de determinismo, y que la hipótesis de Strawson de que ninguna teoría como la determinista podría ser nunca capaz de hacernos adoptar universalmente la actitud objetiva, sólo es cierta si se maneja un concepto de determinismo muy vago según el cual toda acción humana estaría sujeta a una ley. Sin embargo la discusión se hace más complicada si se toman en consideración formas más complejas de entender el determinismo. Ayer propone considerar la hipótesis de que todas las acciones humanas estuvieran sujetas a “leyes manejables; esto es, las generalizaciones deberían ser tales que pudiéramos realmente usarlas para explicar cada faceta del comportamiento humano, y esto no meramente *ex post facto*; deberían capacitarnos para hacer

⁴⁶⁰ Calderón de la Barca, *La vida es sueño*, en Calderón de la Barca 1973: 177.

⁴⁶¹ Strawson, 1962, en Fischer & Ravizza 1993: 60.

⁴⁶² Ayer 1980, en Van Straaten 1980: 4.

consistentemente predicciones con éxito”.⁴⁶³ Una ley manejable viene a ser una ley predictiva que, “si la tuviéramos, sería cuantitativa y usaría conceptos muy distintos” de los de una “generalización como «los cristales de las ventanas son frágiles, y las cosas frágiles tienden a romperse cuando se las golpea con suficiente fuerza, en condiciones normales»”.⁴⁶⁴ El hecho es que precisamente los casos en que normalmente estamos dispuestos a eximir de responsabilidad o a excusar a un agente son aquellos en que contamos con una descripción “científica” suficientemente detallada de los pasos que llevaron a dicho agente a llevar a cabo una determinada acción. El que esto sea así es una buena razón para sospechar que, si fuésemos capaces de contar con descripciones científicas detalladas de nuestros cursos de acción, o incluso de hacer predicciones con éxito, entonces tenderíamos a vernos a todos los seres humanos como eximidos permanentemente de responsabilidad, esto es, estaríamos siempre en la actitud objetiva.⁴⁶⁵

La idea es pues la siguiente: la afirmación de Strawson de que es imposible adoptar de forma permanente la actitud objetiva se basa en el hecho de que nadie ha sido capaz de hacerlo hasta ahora, pero nada autoriza a Strawson a afirmar que nunca nadie será capaz de hacerlo, pues cabe pensar que si contásemos con una teoría científica suficientemente detallada que pudiésemos utilizar de forma efectiva para dar cuenta del comportamiento humano, entonces seguramente adoptaríamos de forma permanente la actitud objetiva. Dicho de otro modo: el argumento naturalista de Strawson es una falacia, pues nada autoriza a deducir del hecho de que hasta hoy nadie haya sido capaz de universalizar la actitud objetiva, el que nunca nadie vaya a ser capaz de hacerlo. Si hasta hoy no lo hemos hecho es porque no contamos con descripciones científicas suficientemente detalladas y sobre todo no contamos

⁴⁶³ Ibidem, p. 6.

⁴⁶⁴ Davidson 1963, en Davidson 1995: 32.

con explicaciones científicas manejables acerca del comportamiento humano. Si hubiese una teoría científica precisa y explicativa que diese cuenta del comportamiento humano podríamos suspender nuestras emociones reactivas.⁴⁶⁶

Por ello no es válida la acusación de inconsistencia que se vierte sobre el incompatibilista que no deja de lado la actitud participante: este hecho es posible porque el incompatibilista no cuenta con una teoría realmente precisa y manejable que le explique qué fue aquello que llevó a un determinado agente en una determinada ocasión a actuar como lo hizo. Un filósofo incompatibilista que cree que una acción dada estuvo determinada físicamente, puede aún así verazmente decir que el agente que llevó a cabo dicha acción pudo haber obrado de otro modo. Se puede pensar que si los fisicalistas incompatibilistas no han renunciado hasta ahora a la actitud participante no es porque sean unos hipócritas o que sean inconsecuentes con sus ideas: sencillamente no cuentan con una teoría que sea operativa a la hora de explicar cómo se da la determinación física de los estados mentales que nos llevan a actuar de una u otra forma.

En cualquier caso falta algo en el argumento naturalista de Strawson, a saber: demostrar que el hecho de que hasta ahora nadie haya sido capaz de eliminar la actitud participante no es contingente sino necesario. La validez de la estrategia naturalista depende de esto, de lo contrario parece una posición dogmática que hace afirmaciones esencialistas de las cuales es incapaz de responder. Así pues, voy a intentar mostrar, mediante un análisis fenomenológico de la responsabilidad, que en efecto el hecho de que nadie hasta ahora haya sido capaz de adoptar de forma permanente la actitud

⁴⁶⁵ En Hernández 2003: 9-10, elaboro un experimento mental, sobre un país llamado Determinilandia, que pretende mostrar a las claras las consecuencias de que esto fuese así.

⁴⁶⁶ Me remito una vez más al experimento mental de Determinilandia como prueba, en Hernández 2003: 9-10.

objetiva es un hecho inevitable. De lograrlo, habría deshecho la falacia de Strawson.

II.3.2. Un experimento fenomenológico

Voy a reconstruir aquí un experimento fenomenológico que lleva a cabo Galen Strawson en su obra *Freedom and Belief*, con el fin de mostrar hasta qué punto cabe afirmar que nuestra actitud participante es natural y que por lo tanto nada podría hacernos dejarla de lado, y mucho menos una cierta convicción de orden teórico.

Galen Strawson señala que “el hecho de que la intuición del incompatibilista [según la cual si el determinismo es cierto entonces no somos responsables] tenga tanto poder sobre nosotros es un hecho tan natural en seres pensantes como nosotros como lo es el hecho de nuestro bastante irreflexivo compromiso con las actitudes reactivas”.⁴⁶⁷ Así pues, como veíamos más arriba de la mano de Ayer, apelando a la naturaleza sin ningún tipo de argumentación filosófica ulterior es imposible demostrar la necesidad de la presencia de la responsabilidad en nuestras vidas ya que, como bien dice Galen Strawson, el incompatibilista puede responder al strawsoniano que su compromiso natural con la duda incompatibilista es tan potente o más que el compromiso que aquel pueda tener con las actitudes reactivas. Por ello es necesario ir más allá en lo que respecta a la investigación de nuestro compromiso con nuestras actitudes reactivas, a fin de descubrir si dicho compromiso es realmente primordial, y por lo tanto debe anteponerse a la duda incompatibilista. Nos hallamos pues ante “un verdadero conflicto de compromisos”⁴⁶⁸ en el que hay que dilucidar qué compromiso, entre los dos que se enfrentan y de los cuales se puede decir que uno es un compromiso

⁴⁶⁷ Strawson, G. 1986: 88.

práctico (el de Strawson) y el otro teórico (el del incompatibilista), es más básico. Para elucidar esta cuestión vamos a sumergirnos, de la mano de Galen Strawson, en un experimento mental consistente en “la rigurosa aplicación de la creencia en el determinismo al curso presente de la propia vida”⁴⁶⁹. Dicho de otro modo: experimentemos seriamente el hecho de creer que estamos absolutamente determinados.

Según Galen Strawson este experimento “debería tener el efecto de eliminar cualquier forma de presencia de un ‘yo’ que decida y actúe libremente (...) pues –parece ser- sencillamente no hay ningún papel que pueda jugar un ‘yo’ así”.⁴⁷⁰ Se pierde el yo en tanto éste se nos revela directamente ligado a la intuición de que nuestras acciones dependen plenamente de él, como ya en su día escribía Schopenhauer: “Existe otro dato de la conciencia (...) se trata del sentimiento, completamente claro y seguro, de la *responsabilidad* por aquello que hacemos, de la *imputabilidad* de nuestras acciones, lo que descansa sobre la certeza inmovible de que somos nosotros mismos los *autores de nuestros actos*.”⁴⁷¹ Así pues, si intentamos ver cualquiera de nuestras acciones como determinada sin ser uno mismo causa de ellas, se esfuma nuestra subjetividad. No debería sernos posible no ya decir “yo he hecho x”, sino ni tan siquiera decir “yo”. Describiríamos cada uno de nuestros movimientos de forma impersonal: no diríamos “yo voy hacia x” sino más bien algo así como “esto va hacia x”. Naturalmente el que esto sea posible, por lo menos de forma continuada, es difícil de creer, pues parece que nuestro compromiso con el “yo” como sujeto de unas acciones que difícilmente podemos dejar de ver como propias sí es un compromiso fundamental.⁴⁷² En cualquier caso, lo que

⁴⁶⁸ Ibidem: 89.

⁴⁶⁹ Ibidem: 96.

⁴⁷⁰ Ibidem: 96-97.

⁴⁷¹ Schopenhauer 2000: 151.

⁴⁷² Sobre este punto habría mucho que discutir, es un problema central de la filosofía de la mente el de la autoridad de la primera persona y el del sujeto lingüístico. Lamentablemente no puedo

permite ver el experimento mental de Galen Strawson es que una concepción determinista fuerte interpretada en términos incompatibilistas no da como resultado un sujeto que encuentra *a posteriori* una causa que determina necesariamente cada una de sus acciones, sino que más bien redundaría en una ausencia de sujeto, en la percepción de uno mismo como un medio en el que se disuelve lo que se suele entender por sujeto. Una vez sumergidos en el experimento mental “ningún pensamiento sería nunca capaz de emerger como producto real del familiar ‘yo’, el putativo originador real de nuestros pensamientos, nuestras decisiones y nuestras acciones que no es meramente un fenómeno determinado: ese ‘yo’ se desvanecería a perpetuidad, sin importar lo lejos que uno persiguiese la posible regresión de pensamientos sobre pensamientos sobre pensamientos”.⁴⁷³

Bien, este experimento es posible y sus consecuencias son las que se acaban de ver, pero considérese ahora este experimento no como tal experimento sino como la forma de vida a la que se ve abocado el incompatibilista que piensa que el determinismo es cierto. Entonces las consecuencias variarían pues no parece que nadie pueda ser capaz de renunciar efectivamente a su “yo” a perpetuidad. Galen Strawson piensa que el incompatibilista, antes de actuar, será plenamente consciente de que “él (esto es, él el verdadero agente responsable, él tal y como automáticamente se concibe a sí mismo en su natural, irreconstruido, pensamiento acerca de sí mismo) no puede hacer nada de nada”.⁴⁷⁴ El incompatibilista probablemente conseguirá verse a sí mismo exento de responsabilidad, como un objeto inadecuado para la aprobación o la desaprobación moral. Sentirá que, aunque no puede liberarse de su sensación de que lo que él hace lo hace él “ya no tiene sentido la idea de llevar a cabo una acción que sea verdaderamente *su* acción.

ahondar en él aquí, adopto pues la postura más comúnmente aceptada que es la que adopta Galen Strawson en su libro.

⁴⁷³ Strawson, G. 1986: 98-99.

Pues la sensación de sí mismo que naturalmente tiene (...) es irremediablemente incompatible con toda aceptación profunda de la idea de que todo lo que es y hace está determinado.”⁴⁷⁵ Así, contrariamente a lo que uno podría pensar, el incompatibilista no decidirá dejar de lado sus actitudes reactivas, pues se dará cuenta de que sencillamente no puede dejar nada, no puede en absoluto elegir qué hacer o qué no hacer. Es más, pensará que esa deliberación que está llevando a cabo acerca del camino a seguir tampoco es realmente suya, que no controla sus pensamientos. De este modo razonar para obrar en consecuencia pierde su sentido. “Así que todo el cuadro del meditabundo incompatibilista determinista que acaba creyendo en el determinismo, y que entonces aborda la pregunta de qué hacer con ello, debe estar mal concebido. La pregunta debe ser completamente irreal para él, siempre que concentradamente se aplique el pensamiento del determinismo a sí mismo. Pues debe sentir entonces que no puede realmente elegir hacer, o hacer, nada, en la forma en que lo pensó.”⁴⁷⁶ Esta es la tragedia del incompatibilista: una vez llega a la situación de conocimiento a la que quería llegar se da cuenta de que no puede obrar en consecuencia porque sencillamente ya no tiene sentido obrar, es más, no tiene sentido sacar ninguna conclusión. “Uno no puede decidir no decidir nada apoyándose en que uno no puede decidir nada.”⁴⁷⁷ En cierto sentido al incompatibilista todo le sucede sin más, él no puede tener ningún pensamiento que conlleve cierto propósito. Debería sentirse como una especie de pantalla en la que se van proyectando imágenes una tras otra y sobre las cuales no tiene ningún poder, y sobre todo de las cuales no puede extraer ninguna conclusión que le permita hacer algo. Está paralizado, no tiene nada que hacer: “(...) no puede decidir abandonarse a su determinación, pues esto también sería algo determinado, y por

⁴⁷⁴ Ibidem: 100.

⁴⁷⁵ Ibidem: 101.

⁴⁷⁶ Ibidem: 102.

consiguiente no algo que realmente hiciera, en el sentido esencial. No puede pensar ‘encuentro que se me presentan las razones para hacer X, y *puesto que* X ahora se me presenta (determinadamente, lo sé) como lo mejor que se puede hacer, haré X’, como si tuviese acceso a una razón adicional para hacer X (...).’⁴⁷⁸

Bien, éstas parecen ser las características de una creencia genuina en el determinismo. Voy a observar a continuación si permiten sacar una conclusión acerca de la importancia de nuestro compromiso con las actitudes reactivas, y por ende averiguar si la estrategia naturalista es un buen método para combatir el incompatibilismo.

II.3.3. Limitaciones del naturalismo

Las consecuencias de mantener una tesis incompatibilista y llegar a creer que el determinismo es cierto serían las que se acaban de ver: no sólo la pérdida de las actitudes reactivas y de cualquier tipo de comportamiento que implique aprobación o desaprobación moral, sino directamente la disolución del sujeto tal y como lo solemos entender o sentir. No obstante, “una persona puede *teóricamente* aceptar completamente que él, o ella, es plenamente un producto de su herencia o su entorno –muchos de nosotros lo hacemos- y aún así *no* tener una concepción como la que aquí se requiere para el genuino incompatibilista determinista (no-auto-determinista). De hecho una auto-concepción tal parece remotamente posible para los seres humanos. Parece requerir la disolución de cualquier sentido reconocible del ‘yo’ humano. Ciertamente uno no puede adoptar una actitud ‘objetiva’ tan radical hacia uno mismo a voluntad.”⁴⁷⁹

⁴⁷⁷ Ibidem.

⁴⁷⁸ Ibidem: 141.

⁴⁷⁹ Ibidem: 99-100.

Así pues, parecería que el experimento fenomenológico del apartado anterior viene a confirmar la sospecha de Peter Strawson de que nuestro compromiso con las actitudes reactivas es tan básico que ninguna exigencia de orden teórico podría nunca ser capaz de hacérselo abandonar. Dicho de otro modo: nuestras actitudes reactivas, nuestro comportamiento como seres responsables, son naturales en el ser humano y por lo tanto nunca podremos ser capaces de dejarlos de lado. La disolución de nuestra voluntad, de nuestra capacidad de decisión, implicaría nuestra destrucción como individuos y es imposible que podamos vivir de tal modo. Así, somos humanos en tanto somos un “yo” libre y por ello no podremos jamás, independientemente de lo que diga la metafísica, dejar de lado nuestra autopercepción de nosotros mismos como seres plenamente responsables.

Naturalmente uno podría decir que Galen Strawson exagera las consecuencias que tendría el que un incompatibilista viniese a descubrir que el determinismo es cierto. Él mismo es consciente de esta posible crítica y señala que “uno podría sencillamente dejar de creer que concretamente las actitudes reactivas morales y las prácticas de elogio y culpa estaban justificadas o eran apropiadas, y perder esta creencia podría causar el que uno dejase de estar movido por el elogio y la culpa. Generalmente, creer que el determinismo es cierto y ser (...) muy racionalístico podría llevar a un hombre a tener más actitudes objetivas, sin que él intentase *hacer* nada. [...] De otra parte, sus actitudes morales reactivas y no-reativas podrían no resultar afectadas por su nueva creencia teórica, dada la fuerza de su compromiso con ellas.”⁴⁸⁰

Bien, esta podría ser también la consecuencia de ser incompatibilista y venir a creer que el determinismo es cierto. No obstante, lo que se puede deducir de esto no es diferente de lo que ya se había dicho: Peter Strawson parece tener razón, nada puede hacernos abandonar para siempre nuestras

⁴⁸⁰ Ibidem: 103

emociones reactivas. O bien porque hacerlo implicaría la disolución de nuestro yo, o bien sencillamente porque nuestro compromiso con la coherencia teórica es muy inferior a nuestro compromiso con nuestra vida como seres responsables. En fin, el experimento fenomenológico llevado a cabo parece confirmar que nuestro compromiso con las actitudes reactivas no es contingente sino absolutamente necesario, y por lo tanto forma parte de nuestra naturaleza. No hay ninguna falacia en el argumento naturalista de Strawson. ¿O sí?

Las consecuencias teóricas del experimento fenomenológico son claras; no obstante debería analizarse con más detalle ese experimento partiendo de otros supuestos. No olvidemos la objeción que le hacía Ayer a Strawson: imaginémonos otro mundo en el que contásemos con leyes “manejables” que nos permitiesen hacer predicciones y explicar con detalle el comportamiento humano. A continuación hagamos el experimento mental de Galen Strawson con los habitantes de ese mundo.⁴⁸¹ Parece que no les costaría tanto renunciar a ese “yo” libre que antes parecía tan inamovible. En efecto, ellos serían capaces de ir dando cuenta de su vida de forma muy precisa sin necesidad de recurrir a la idea de voluntad y seguramente sin actitudes reactivas. Y en cualquier caso pensar en ello no es autocontradictorio.

Así pues lo que cabe decir es que probablemente abandonar plenamente nuestras actitudes reactivas es imposible para nosotros hoy por hoy en tanto sujetos que pertenecemos a una cultura concreta y que nos hallamos en un momento histórico concreto. Pero cabe pensar en una evolución de nuestra cultura hacia otra en que el concepto de sujeto fuese totalmente diferente y totalmente compatible con la ausencia de libertad tal y como la entendemos hoy en día. Las consecuencias del experimento fenomenológico de Galen Strawson no son universales y no nos dicen nada

⁴⁸¹ Véanse las notas 465 y 466.

acerca de la naturaleza humana, no responden en absoluto a la crítica de Ayer. En el fondo, aunque lo que pretende Ayer es dar con una fórmula de determinismo más interesante para el debate compatibilismo vs. incompatibilismo, lo que realmente lleva a cabo es imaginar un mundo posible en que la tesis del determinismo fuese algo más que una hipótesis metafísica, esto es, en que fuese una ley científica antropológica operativa. Así, de la mano de Ayer se puede imaginar otra sociedad en que el determinismo fuera funcional socialmente, y frente a este otro experimento mental no vale la prueba fenomenológica de Galen Strawson, ésta sólo es aplicable a nosotros aquí y ahora; o cuando menos puede decirse que sólo nosotros somos capaces de extraer de ella las consecuencias terribles que el determinismo y el incompatibilismo parecen tener para el sujeto. Pero cabe imaginar otras sociedades en que lo que nosotros juzgamos insostenible estuviese a la orden del día. Tal vez una sociedad como la que postula Ayer, dominada por leyes deterministas “manejables” socialmente, no le encontrase ningún sentido a plantearse el problema de la disolución de su yo ante la falta de una libertad contra-causal.⁴⁸²

A lo largo de la historia, ¿a cuántas ideas que parecían inamovibles no ha renunciado ya el ser humano? Se puede pensar que la idea actual del sujeto libre no es más importante para nosotros de lo que lo era el que nuestro planeta fuese el centro del universo para los precopernicanos. Supóngase con Ayer que llegamos a conseguir leyes manejables que explican paso a paso el comportamiento humano: usando la terminología kuhniana, probablemente acabaría surgiendo un cambio de paradigma y obtendríamos un concepto de sujeto libre totalmente diferente que nos permitiría fácilmente abandonar nuestras actitudes reactivas y empezar a proceder únicamente de manera “objetiva”. Sin ir más lejos Galen Strawson suministra en *Freedom and Belief* un

⁴⁸² Idem nota anterior.

ejemplo que explica muy bien lo que estoy tratando de mostrar: el de una cultura⁴⁸³ que realmente consigue o intenta llevar a cabo esa disolución del sujeto libre que se ha planteado como la consecuencia necesaria de una creencia genuina en el determinismo. Existen pensadores budistas que opinan que la idea de un yo individual es una ilusión⁴⁸⁴ y no sólo eso, sino que además llevan a cabo prácticas de meditación destinadas a perder esa percepción del yo que para ellos es un engaño. Así, “alguien que tuviese tal creencia [que fuese incompatible y pensase que el determinismo es cierto], y deseara perder cualquier sentido del sujeto como libre o en cualquier caso verdaderamente auto-determinado, en orden a llevar a cabo una actitud más correcta hacia el mundo, haría bien en adoptar la susodicha práctica de la meditación que elimina el ‘yo’.”⁴⁸⁵

Así, Ayer parece tener razón: nada justifica la tajante afirmación de Strawson según la cual nos es imposible renunciar a nuestras actitudes reactivas. “No hay duda de que la decisión de adoptar la actitud ‘objetiva’ hacia uno mismo y los demás no puede ser implementada de la noche a la mañana, dada la naturaleza y la fuerza del entramado y nuestro compromiso con él. [...] Pero si admitimos la posibilidad de alteraciones parciales en actitudes o hábitos de pensamiento con los cuales estamos profundamente comprometidos, entonces esto apunta la posibilidad de un progresivo abandono de esas actitudes y hábitos de pensamiento que, llevado a término gradualmente, asciende a un abandono total en lo relativo a nuestra posición original.”⁴⁸⁶ Nada en nuestra naturaleza parece implicar la irrenunciabilidad a la libertad, se trata más bien de un rasgo cultural muy profundamente arraigado. La estrategia naturalista es pues estéril *por sí sola* para justificar nuestras prácticas

⁴⁸³ Utilizo aquí la palabra “cultura” de forma muy vaga: me refiero a un grupo de personas que comparten un cierto paradigma.

⁴⁸⁴ Se trata de la doctrina del *satkayadrsti*, “la falsa visión de la individualidad”.

⁴⁸⁵ Strawson, G. 1986: 119. Hay que señalar que Galen Strawson hace esta observación con cierto tinte irónico.

de atribución de responsabilidad. Esto no quiere decir que el experimento fenomenológico de Galen Strawson que he reconstruido aquí no tenga interés. Sólo que es falible, o mejor dicho: que los datos de la autoconciencia por sí solos no son suficientes para demostrar un compromiso irrenunciable con las actitudes reactivas, pues si nuestra autoconciencia fuese formada en un contexto cultural diferente, ésta podría revelar compromisos totalmente diferentes. Dice Schopenhauer: “La dependencia de nuestro hacer, esto es, de nuestras acciones corporales, con respecto a la voluntad, que es lo que nos testimonia la autoconciencia, es algo muy distinto de la independencia eventual de nuestros actos de voluntad con respecto a las circunstancias exteriores (...).”⁴⁸⁷ Así, unas leyes manejables podrían ser unas “circunstancias exteriores” apropiadas para dejar de lado las actitudes reactivas.

En cualquier caso, la investigación fenomenológica, al mostrar lo íntimamente ligados que están nuestro compromiso con la responsabilidad y las actitudes reactivas, y nuestro compromiso con el sujeto, nos revela que difícilmente podría prescindirse de las actitudes reactivas y de la responsabilidad. Este hecho es sin duda importante, pero no puede constituir un argumento en contra de la posición incompatibilista (el incompatibilista no niega que lo que está en juego sea extremadamente importante), tan sólo puede comprenderse como una iluminativa explicitación de nuestra condición moral actual, si se toma como un argumento (sobre todo definitivo como pretende Strawson) se corre el riesgo de caer fácilmente en una falacia o, aún peor, en un esencialismo para siempre impugnado por un incompatibilista muy en su derecho de no creer en *a priori* metafísicos.

Así pues, sería conveniente impregnar la defensa de las actitudes reactivas de una buena dosis de falibilismo, y argumentar más bien en pro de la racionalidad de nuestro funcionamiento como seres responsables que en pro

⁴⁸⁶ Ibidem.

⁴⁸⁷ Schopenhauer 2000: 60.

de su indisoluble vinculación con nuestra existencia. Esto es lo que intentaré hacer en el siguiente apartado reinterpretando desde un pragmatismo de corte wittgensteiniano la experiencia fenomenológica de nuestro yo como aglomerado de nuestros deseos y razones. El carácter está de vuelta.

II-4. El yo y su heteroconciencia: vivir en un mundo participante

“Mi padre, que está presente, / por excusarse a la saña / de mi condición, me hizo / una fiera humana: / de suerte, que cuando yo / por mi nobleza gallarda, / por mi sangre generosa, / por mi condición bizarra / hubiera nacido dócil / y humilde, sólo bastara / tal género de vivir, / tal linaje de crianza, / a hacer fieras mis costumbres (...).”⁴⁸⁸

Más arriba señalé de la mano de Paul Russell que el principal problema con el que se encuentra la estrategia racionalista es el error que comete Strawson al equiparar incapacidad y anormalidad. Pues bien, hay dos formas de leer este error.

Una de ellas es hacer una lectura esencialista de su argumento normativista; lo que da como resultado que por naturaleza el ser humano es capaz de enfrascarse con los demás en relaciones interpersonales normales, y por ello en cierto sentido ser incapaz de hacerlo es una anormalidad y además no cabe imaginar que dejará de serlo. Si por naturaleza ser capaz es ser normal, entonces por naturaleza ser incapaz es ser anormal. Bien, esta lectura del argumento normativista lo convierte en una nueva versión del argumento naturalista en sentido fuerte, y por lo tanto se ve sujeta a la crítica que he hecho de este argumento en el apartado anterior. No sería raro pensar que esta lectura de Strawson es la correcta, pues sería una lectura de este tipo la que permitiría comprender que Strawson opte por desarrollar una argumentación que, como ha sido mostrado más arriba, es incoherente en tanto la estrategia

⁴⁸⁸ Calderón de la Barca, *La vida es sueño*, en Calderón de la Barca 1973: 276.

naturalista y la racionalista no se pueden solapar (aunque, como se ha visto, puedan ser complementarias si se matizan los argumentos naturalistas).

Sin embargo, se puede hacer una lectura más rica del argumento normativista de Strawson, conciliable con su argumento pragmatista y que nos solucionaría de entrada buena parte de los problemas que lleva consigo la teoría de Strawson y a los que habrá ocasión de enfrentarse más adelante. Esta nueva lectura sería como sigue: hay una forma normal de relacionarse con los demás en nuestras sociedades. Poco a poco esa forma ha ido imponiéndose como una suerte de profundo contrato social implícito hasta hacernos a todos miembros de una comunidad moral cuyos límites se medirían gracias a nuestra capacidad para, mediante las actitudes reactivas y (y esto es importante subrayarlo) mediante la suspensión de las mismas en los momentos adecuados, enfrascarnos en relaciones interpersonales normales. Hay pues que entender la normalidad como adaptación a una pauta social base, muy bien establecida, marca de nuestra comunidad moral.⁴⁸⁹ Nada dice que no podamos hacer algún día que la incapacidad aparezca como lo normal, pero entonces sería necesario un buen argumento para hacerlo y el incompatibilismo no es ese argumento.

Así es cómo se unen el argumento normativista y el pragmatista: el primero dice cómo están las cosas, el segundo dice que lo mejor es que sigan como están.⁴⁹⁰ Ambos argumentos son absolutamente necesarios y por ello también, desde el incompatibilismo, es necesario un argumento que demuestre que sería mejor comportarnos como si nouviésemos responsabilidad moral. Pero los incompatibilistas no dan nunca ese argumento, tan sólo dicen que lo otro no sería racional. Y sin embargo, como dice Strawson “no sería necesariamente racional elegir ser más puramente racionales de lo que

⁴⁸⁹ El desarrollo y justificación de esta idea es lo que me ocupará en la tercera y última parte del trabajo.

⁴⁹⁰ Y se puede tener en cuenta el argumento naturalista como un argumento normativista más: nuestro compromiso con las actitudes reactivas es lo natural *hoy por hoy*, luego *hoy por hoy* “yo” y “libertad” están indisolublemente unidos.

somos”.⁴⁹¹ Sin ir más lejos, todos sabemos desde que nos lo reveló Hume que la inducción no es un método válido de conocimiento, pero no por ello hemos renunciado a ella, pues considerando las ganancias y las pérdidas que conllevaría su desaparición como principio de conocimiento preferimos dejar las cosas como están.⁴⁹²

Así pues, hay una forma de desarrollar la estrategia racionalista de Strawson interpretando su argumento normativista en clave pragmatista o, si se quiere, desarrollando una versión pragmatista del racionalismo. Esta versión de la estrategia racionalista es la que voy a desarrollar a continuación y gracias a ella trataré de mostrar hasta qué punto la visión incompatibilista de la libertad es incorrecta, con lo que, espero, habré dado por fin con la justificación última de la responsabilidad moral.

II.4.1. Una nueva versión de la estrategia racionalista

Voy a tratar de mostrar ahora que las intuiciones de Strawson son, una vez más, correctas. Sin embargo la forma en que dichas intuiciones aparecen en su artículo es inadecuada, así que aquí voy a intentar darles una forma más interesante. Para empezar hay que desterrar, por fin, ciertos rasgos terminológicos que no dejan comprender correctamente la cuestión que nos ocupa.

En primer lugar no es lícito distinguir entre actitud participante y actitud objetiva sino más bien entre dos tipos diferentes de actitud dentro de un mismo tipo de vida: el participante. En efecto, la actitud en que tienen lugar

⁴⁹¹ Strawson 1962, nota 4, en Fischer & Ravizza 1993: 56.

⁴⁹² Esto no quiere decir que la crítica de Hume a la inducción haya quedado sin consecuencias, pero éstas no ha sido tan radicales como cabría esperar. Del mismo modo nuestra comprensión de que la libertad humana está determinada seguramente lleva consigo ciertas consecuencias que deberían suponer cambios en nuestra forma de entender la libertad, pero nunca consecuencias tan radicales como las que pretendería extraer de ello el incompatibilista.

las actitudes reactivas no es, o no tendría por qué ser, menos objetiva que la actitud que conllevaría la suspensión de dichas emociones. Así, cuando yo me siento indignado porque, siendo dueño de una joyería, me han robado mi pulsera de oro más preciosa, no estoy en una actitud menos objetiva que aquella en la que me encontraría de saber que aquello fue fruto de la compulsión a robar de alguien que sufría de cleptomanía. Strawson maneja una idea de objetividad demasiado restringida, cercana a esa “visión desde ninguna parte” de Nagel,⁴⁹³ que le obliga a pensar que, si estamos llevando a cabo el juicio sobre expectativas que conlleva una actitud reactiva, entonces nunca estamos siendo objetivos. Este error reposa en un prejuicio típicamente filosófico según el cual, si algo pertenece al terreno de lo emotivo, entonces es absolutamente irracional y por ende absolutamente carente de objetividad. Pero esto no tiene por qué ser así y por ello aquí voy a redefinir y a renombrar ciertas cosas.

No existe una actitud participante que se oponga a una actitud objetiva: siempre estamos participando en relaciones interpersonales normales. Cuando adoptamos lo que hasta ahora he llamado una “actitud objetiva”, no estamos fuera de dichas relaciones, sino siguiendo sus reglas que nos indican que, en ciertos casos, debemos suspender nuestras actitudes reactivas. Por ello voy a considerar que la participación con los demás en relaciones interpersonales es algo más amplio, es en general nuestra posición social. Con Habermas se puede decir que se trata de “condiciones presuntamente universales, pero sólo irrebasables *de ipso*, que deben estar satisfechas para que puedan producirse determinadas prácticas para las cuales no hay equivalentes, funcionales imaginables dentro de nuestras formas de vida socioculturales”⁴⁹⁴, esto es, las relaciones interpersonales se dan “sobre el trasfondo de un mundo de la vida

⁴⁹³ Véase Nagel, Thomas, 1986, *The View From Nowhere*, New York: Oxford University Press.

⁴⁹⁴ Habermas 2002: 19.

intersubjetivamente compartido”.⁴⁹⁵ En términos wittgensteinianos, participar con los demás en relaciones interpersonales es nuestra “forma de vida”, nuestra “imagen del mundo”. Así pues voy a hablar de un *mundo participante*.

Bien, dentro de ese mundo participante podemos, a la hora de relacionarnos con los demás, adoptar dos actitudes diferentes: en una de ellas (a partir de aquí *actitud reactiva*) reaccionamos mediante emociones ante ciertas acciones llevadas a cabo por los demás o por nosotros mismos; en la otra (a partir de aquí *actitud reflexiva*) juzgamos acciones intentando suspender nuestras emociones, y analizamos con detenimiento la causas que obligaron a cierto agente o a nosotros mismos a llevar a cabo una determinada acción.

¿Qué se gana con esta redefinición?

En primer lugar se amplían los límites de la racionalidad hasta cubrir con ella todo nuestro mundo participante, con su parte reactiva y su parte reflexiva, y mediante ello tanto las atribuciones de responsabilidad como las exculpaciones. De este modo se puede comprender que ese compromiso tan profundo en el que insisten reiteradamente tanto P. F. Strawson como Galen Strawson no es con una idea, *sino con una forma de vida*. “Para el comportamiento guiado por reglas, las reglas constitutivas abren continuamente la alternativa entre seguimiento de la regla y vulneración de la regla. Además, por principio existe siempre la alternativa entre tener capacidad para hacer algo o no tenerla. Quien no domina la regla de un juego y, por tanto, nunca puede cometer un error, no es un jugador. Esto se muestra *en el curso* de la práctica.”⁴⁹⁶ Por ello el argumento necesario para hacernos abandonar nuestro mundo participante no puede llevar consigo una exigencia de orden teórico, sino práctico. Es necesario demostrar que hay otra forma de vida que sería más provechosa para hacernos abandonar aquella en que

⁴⁹⁵ Habermas 2000: 40. Me resulta especialmente interesante que Habermas recurra al husserliano “mundo de la vida”, que considero equiparable al “mundo participante” al que le dedico la tercera parte del trabajo.

estamos inmersos ahora. Es necesario demostrar que un mundo regido por leyes deterministas manejables sería preferible a nuestro mundo participante en que nos dejamos guiar por la idea de responsabilidad. La metafísica incompatibilista es estéril porque no da cuenta de esta demanda práctica.

En segundo lugar, con esta reelaboración de la estrategia racionalista desaparece (puesto que se elimina el concepto de “actitud objetiva”) un error de Strawson según el cual nuestra actitud sólo podría ser objetiva de no haber emociones de por medio. Gracias a esto se puede rescatar su intuición de que existen emociones esencialmente morales y resolver una contradicción señalada más arriba⁴⁹⁷ (y a la que habrá ocasión de volver más abajo), pues se advierte que suspender nuestras emociones reactivas no viene dado forzosamente porque consideremos a un determinado agente como incapacitado para vivir en nuestro mundo participante. Al contrario: sólo si consideramos que alguien sería susceptible de pertenecer a nuestro mundo participante tiene sentido que adoptemos para con él la actitud reflexiva. No tiene sentido en cambio adoptar dicha actitud para con un animal, porque esto supondría atribuirle la posibilidad de haber obrado responsablemente. El que juzguemos adecuado mirar no-reactivamente a un sujeto sólo cabe en el caso de que seamos capaces de ver a ese sujeto como un candidato posible a haber sido juzgado responsable. Por ello se puede, como defendía Strawson, llevar a cabo esa suspensión de nuestras emociones reactivas también frente a personas sanas y maduras “como un refugio de las tensiones de la implicación; o como una ayuda para la política; o simplemente por curiosidad intelectual”.⁴⁹⁸ Porque estar en el mundo participante es una “suposición constitutiva” (usando una vez más el vocabulario de Habermas), un fundamento práctico primordial, mucho más básico que cualquier

⁴⁹⁶ Habermas 2000: 41.

⁴⁹⁷ Ya señalada en la nota 422.

⁴⁹⁸ Strawson 1962, en Fischer & Ravizza 1993: 52.

compromiso de carácter teórico, por ello tiene sentido que podamos por un lado atribuir responsabilidad y por otro lado exculpar. Recuérdesse lo que decía Paul Russell: “Lo que es particularmente molesto en la estrategia naturalista de Strawson, expresado en términos generales, es que arroja dudas sobre nuestra habilidad o capacidad para contener o controlar nuestra vida emocional de acuerdo con los dictados de la razón.”⁴⁹⁹ Pues bien, esas dudas que arroja la estrategia naturalista de Strawson son las que me llevaron a hablar en la nota 415 de una contradicción en su argumentación (cuyo alcance, insisto, será puesto de relieve más claramente en *III.1.2*). Dejando atrás la estrategia naturalista fuerte, ya no hay lugar para esa contradicción. Puesto que la razón, aún expresada mediante emociones, es la que guía y ordena el mundo participante; siempre es posible, cuando se juzga que hacerlo sería razonable, dejar de lado nuestras emociones reactivas y adoptar la actitud reflexiva. Lo normal será hacerlo ante casos de incapacidad para participar normalmente en nuestro mundo participante, pero no siempre. El del psicoterapeuta y el paciente es el ejemplo más manejado de actitud reflexiva, y su abuso conduce a una visión reduccionista de lo que sería estar en la actitud reflexiva. En cambio tal vez no sería un mal ejemplo el de un político cuando transmite un mensaje electoral o el de un cura que pretende persuadir a un adolescente de que la masturbación es un pecado o el de cualquier agencia de publicidad cuando maquina como conseguir que algo que nunca quisimos se nos haga indispensable. La actitud reflexiva es mucho más rica de lo que la caracterización que de la actitud objetiva hizo Strawson permitiría suponer. Por ello Strawson sacó a la actitud objetiva del mundo participante, por ello incluyo ahora en él la actitud reflexiva como parte indispensable que dota de sentido a la actitud reactiva.

⁴⁹⁹ Russell, P. 1992: 297.

Por último, y este punto es el que voy a desarrollar en el apartado siguiente, esta nueva versión de la estrategia racionalista permite entender mejor dónde se halla el incompatibilista y por qué su postura hoy por hoy es insostenible.

Ahora, para ganar en claridad antes de continuar voy a dejar constancia mediante un gráfico de la reestructuración que de los conceptos de Strawson he realizado. Hay que decir que este gráfico deberá ser completado ulteriormente con los diferentes tipos de emociones que pueblan nuestra actitud reactiva y con diferentes tipos de exenciones y excusas que componen la actitud reflexiva.

Nótese que tanto desde la actitud reflexiva como desde la actitud reactiva es posible atribuir responsabilidad y exculpar. A este respecto me voy a permitir aquí, pues no me será posible hacerlo ya más adelante, un pequeño excursus para precisar en qué consiste hacer exculpaciones emotivas y en qué consiste hacer atribuciones reflexivas de responsabilidad.

Entiendo por exculpaciones emotivas lo que creo se puede llamar *disculpas*. Se trataría de excusas que no son válidas pero que, de hecho, utilizamos y que a veces funcionan. Me estoy refiriendo a expresiones como las siguientes: “sí, pero es tan guapo”, “sí, pero es tan graciosa”, “sí, pero es tan majo”. O tal vez también del estilo de: “es que él es así”, “ya la conoces”, “pero tiene buen fondo”, “ya no va a cambiar”. Considero que las disculpas existen y funcionan, pero tienen un rol marginal en nuestras vidas, por lo menos desde el punto de vista moral. Por ello no me ocuparé más de ellas, aunque no dejaré de tenerlas en cuenta como parte del mundo participante.

Por otra parte, considero que sí son formas genuinas de hacer juicios sobre responsabilidad los *veredictos*, como creo se pueden llamar las atribuciones reflexivas de responsabilidad. Probablemente los veredictos que se dan en los juicios no sean los únicos casos de atribución reflexiva de

responsabilidad, pero sí los considero un caso paradigmático y por eso les doy su nombre al todo.

Muchas cosas han de ser aclaradas todavía para que sea posible leer en toda su riqueza este gráfico, así pues figura aquí de forma provisional para que sea posible visualizar cómo encajan algunos conceptos que he venido estudiando:



II.4.2. Mundo participante e incompatibilismo

Tal y como se pudo ver gracias al experimento mental de Galen Strawson, el incompatibilista determinista se encuentra, en cierto sentido, fuera del mundo. No es tanto que el incompatibilista, como pretendía Strawson, opine que deberíamos universalizar la actitud reflexiva sino, más ampliamente, abandonar el mundo participante (atendiendo a la nueva versión de la vía racionalista que he comenzado a esbozar en el apartado anterior). En efecto, para el incompatibilista esa disolución del sujeto, tal y como éste es entendido normalmente, que conllevaría la creencia de que el determinismo es cierto, sería algo más que adoptar una actitud diferente, más objetiva. Se trataría de

algo mucho más radical, habría que dotar de un nuevo sentido a nuestras acciones, sería otra forma de vida, otro mundo, un absoluto cambio de paradigma. Así lo explica Susan Wolf: “La verdad del determinismo, pues, no nos da ninguna razón en absoluto para reemplazar nuestras actitudes reactivas presentes con la actitud objetiva. Algunos pueden pensar, no obstante, que nos da una razón para hacer algo todavía más drástico –esto es, abandonar para siempre la adopción de actitudes. (...) [S]i el determinismo fuera cierto, y ello implicase, a efectos de un hecho metafísico, que no fuéramos seres libres y responsables, entonces la única forma en que podríamos vivir de acuerdo con los hechos sería abandonando para siempre la adopción de actitudes.”⁵⁰⁰

No es casual que el monje budista que aspira a hacer desaparecer su “yo” abandone la sociedad en que vive y se retire al bosque o a un templo. Una vez hubiese logrado su objetivo le sería imposible comunicarse, no sería capaz de sobrevivir en nuestra sociedad porque realmente habría pasado a vivir en un mundo diferente.⁵⁰¹ La responsabilidad sólo significa algo en nuestro juego de lenguaje actual. Únicamente dentro de él tiene sentido distinguir entre personas capaces e incapaces, y no fuera. A lo que aspira el incompatibilista es a salir de este juego del lenguaje, no a redefinir algunos términos, es a establecer una forma nueva de normalidad, no a que todos seamos vistos como incapacitados para las relaciones interpersonales normales. En este sentido lo que hace el incompatibilista es jugar mal el juego del lenguaje en que estamos, porque dentro del mismo no tiene sentido lo que dice ni lo que pretende.

Nuestra responsabilidad se define, se crea, intersubjetivamente, mediante la interacción con los demás. Hay ocasiones en que esa interacción

⁵⁰⁰ Wolf 1981, en Fischer & Ravizza 1993: 113.

⁵⁰¹ Entiendo aquí “mundo” como una configuración social, como un juego de lenguaje según el sentido que le da Wittgenstein a esta expresión en sus *Investigaciones Filosóficas*, 23: “La expresión ‘juego de lenguaje’ debe poner de relieve aquí que *hablar* el lenguaje forma parte de una

es frustrada, pero esto sucede como excepción y precisamente por serlo somos capaces de considerarlo anormal, porque altera nuestra forma común de comportarnos en sociedad, nuestras “relaciones interpersonales normales”. En este sentido el incompatibilista es una especie de extraterrestre, incapaz de comunicarse correctamente y que emite oraciones sin sentido porque pretende que lo tengan fuera del sentido que les ha dado la sociedad intersubjetivamente, y fuera de ese sentido no hay nada. “(...) [N]o tengo mi imagen del mundo porque me haya convencido a mí mismo de que sea la correcta; ni tampoco porque esté convencido de su corrección. Por el contrario, se trata del trasfondo que me viene dado y sobre el que distingo entre lo verdadero y lo falso.”⁵⁰²

Desde la versión pragmatista del racionalismo de Strawson que he adoptado se puede equiparar nuestra autopercepción como seres libres y responsables con nuestra creencia de que el mundo exterior existe. Así, la responsabilidad pasa a formar parte de nuestra imagen del mundo, algo que nos acompaña en todo momento, algo que somos y sobre todo algo que nos permite actuar. Esto último, si se sigue a Wittgenstein, convierte la responsabilidad en una evidencia pues “una evidencia segura es la que se *supone* incondicionalmente segura, algo de acuerdo con lo que *actuamos* con seguridad, sin duda alguna.”⁵⁰³ Ciertamente, la filosofía lleva siglos dedicándose a desmontar evidencias, “sin embargo, la fundamentación, la justificación de la evidencia tiene un límite; -pero el límite no está en que ciertas proposiciones nos parezcan verdaderas de forma inmediata, como si fuera una especie de *ver* por nuestra parte; por el contrario, es nuestra *actuación* la que yace en el fondo del juego del lenguaje”.⁵⁰⁴ Así, mientras el incompatibilista no sea capaz de actuar

actividad o de una forma de vida”. De hecho a partir de aquí usaré indistintamente ambas expresiones.

⁵⁰² Wittgenstein 1988, § 94: 15c.

⁵⁰³ Ibidem, § 196: 27c.

⁵⁰⁴ Ibidem, § 204: 28c.

conforme al juego del lenguaje que pretende inventar, a la imagen del mundo alternativa que propone, no puede decirse que realmente esa imagen exista más que como un mal uso del lenguaje, como una mala comprensión de la imagen del mundo que maneja. O ni tan siquiera eso: en tanto el incompatibilista actúa responsablemente y hace atribuciones de responsabilidad está demostrando que comprende la imagen del mundo en que somos seres responsables, es más, actuar responsablemente es lo mismo que creer en la responsabilidad y comprenderla: “Nuestro hablar obtiene su sentido del resto de nuestra actuación”.⁵⁰⁵

Siguiendo esta perspectiva pragmatista Michael McKenna dice: “(...) la capacidad para actuar como un agente moralmente responsable puede ser entendida en analogía con el dominio de un hablante competente de un lenguaje natural. [...] [S]i el determinismo amenazase la capacidad para actuar como un agente moralmente responsable, amenazaría también la capacidad del agente para dominar y hacer uso de un lenguaje natural. En tanto parece que el determinismo no supone esa amenaza para el uso competente del lenguaje, de un modo similar podría argumentarse que el determinismo no supone ninguna amenaza para la responsabilidad moral.”⁵⁰⁶ Mientras, mediante nuestra capacidad para obrar en consecuencia, mostremos que comprendemos lo que significa la responsabilidad, el incompatibilismo no será un verdadero problema. La duda incompatibilista tendría sentido si nos inhabilitase para seguir actuando como actuamos normalmente en nuestras relaciones interpersonales diarias, pero no lo hace, y no lo hace porque no tiene sentido en nuestro contexto vital, en nuestro mundo participante, y no se puede formular fuera de él: “La cualidad moral del motivo de un agente requiere que el agente formule ese motivo dentro de un contexto interpretativo. Ese contexto está fijado por un conjunto de convenciones sociales que estructuran

⁵⁰⁵ Ibidem, § 229: 30c.

⁵⁰⁶ McKenna 1998b: 6-7.

y mantienen una asociación de diferentes tipos de acciones y diferentes tipos de motivos. Del mismo modo que la capacidad de un hablante competente para usar un lenguaje para expresar sus propios pensamientos, un agente moralmente competente es capaz de formar sus propios motivos moralmente imponibles sólo si posee un dominio suficiente de las convenciones que asignan interpretaciones de motivos-tipo a acciones-tipo”⁵⁰⁷

Ser incompatibilista sería, pues, en el mejor de los casos, un juego especulativo intelectual; en el peor, una falta de dominio de las convenciones que nos permiten relacionarnos normalmente con los demás en nuestro mundo participante. Normalmente los incompatibilistas suelen ser personas normales y bien integradas en la sociedad (en la medida en que puede estarlo un filósofo), así pues debe entenderse el incompatibilismo como una hipótesis de otro mundo posible, como un experimento mental. Y como tal no debe preocupar, pues no amenaza nuestra libertad ya que: si tiene sentido, entonces se asienta en una imagen del mundo en que la libertad es un trasfondo imprescindible; y si pretende plantearse desde fuera de esa imagen del mundo, que es la que todos compartimos, entonces no tiene sentido. “Una duda sin término no es ni siquiera una duda.”⁵⁰⁸ Así es como debe entenderse la indisoluble unidad entre nuestro yo y su libertad, que preside nuestra autoconciencia, así es como se justifica el resultado del experimento fenomenológico del apartado anterior. Ese dato de la *autoconciencia*, inamovible aparentemente una vez aparece, esa supuestamente perfecta interdependencia del yo y sus acciones, se basa en un juego del lenguaje, en una imagen del mundo, en fin, en la *heteroconciencia*. *La autoconciencia de la responsabilidad es fruto del mundo participante en que vivimos.*

Nuestra propia autopercepción como seres responsables es imposible sin la presencia del otro desde el principio, esto es, la responsabilidad sólo

⁵⁰⁷ Ibidem.

⁵⁰⁸ Wittgenstein 1988, § 625: 83c.

tiene sentido en la vida social, no como absoluto metafísico. Retomando la cuestión que me ocupó en la parte anterior, llegar a atribuir responsabilidad basándose en las posibilidades alternativas es imposible sin el otro: la idea de que uno “podría obrar de otro modo” surge, como ya se vio, *a posteriori*, es fruto de la autopercepción de nuestra sustituibilidad en la acción. Ello explica el uso educativo de expresiones como “mira a Juanito como él *sí* obedece” o la archifamosa “¿y si Juanito se tira por la ventana, tú también te tiras?”. Estas frases nos muestran, cuando niños, que otro, en nuestro lugar, podría hacer otra cosa. Especialmente la segunda nos enseña que también uno mismo puede elegir qué hacer, que *todos*, los unos con respecto a los otros, podríamos hacer otra cosa y por ello somos libres y responsables. ¿Cómo podría llegarse a esta conclusión sin el concurso del otro en nuestras vidas? De ninguna forma, porque contar con posibilidades alternativas es saber que “*una o más* personas pueden actuar de forma diferente en situaciones similares”. Si creciésemos solos en el mundo, seguramente nos autoperibiríamos como absolutamente necesarios. La responsabilidad presupone un otro que me va a juzgar porque tiene puestas ciertas expectativas en mí, dado que sabe que, de estar en mi situación, él sería capaz de cumplirlas. La responsabilidad nace pues de la heteroconciencia, de la existencia misma del mundo participante como forma de vida.

II.4.3. Las razones de la vida participante

Después de un largo periplo en que he discutido de forma pormenorizada las razones que permiten justificar la permanencia del concepto de responsabilidad moral en nuestras vidas, se puede concluir que la principal crítica que se le puede hacer a Strawson es que cabe imaginar otro mundo en que las leyes científicas nos darían la posibilidad de prescindir de las

emociones reactivas, y por lo tanto seríamos capaces de prescindir de la idea de responsabilidad moral. Esta crítica lleva consigo una certeza: la de que no puede darse una solución *a priori* al problema de la responsabilidad moral (por ello el debate metafísico de la primera parte de este trabajo se llevó a cabo fenomenológicamente, analizando, *a posteriori*, la libertad de la voluntad *como hecho*). Dicho problema debe resolverse mediante un debate pragmatista sobre la racionalidad de la elección de la continuidad de nuestro mundo participante o de otro mundo alternativo, sólo vagamente esbozado por el incompatibilista. De este modo se alcanza la idea de que es indispensable contar con un criterio que permita decidir racionalmente qué vida es preferible: si una dominada por las actitudes reactivas u otra dominada por leyes científicas manejables.

La idea de Strawson es que deberíamos “escoger racionalmente sólo a la luz de una valoración de las ganancias y pérdidas para la vida humana, su enriquecimiento o empobrecimiento”⁵⁰⁹, y que siguiendo este criterio de racionalidad deberíamos mantener nuestra idea común de la responsabilidad moral y continuar moviéndonos a base de actitudes reactivas tal y como venimos haciendo hasta ahora. Sin embargo, Ayer subraya que existe otra visión de la racionalidad, presente incluso en el artículo de Strawson, según la cual la racionalidad de nuestro comportamiento no dependería de ganancias o pérdidas para nuestra vida, sino de su vinculación con ciertas creencias que juzgásemos racionales. Dicho de otro modo: se puede pensar que, si llegásemos a creer que el concepto de responsabilidad moral es irracional, yuviésemos la posibilidad de renunciar a él y a los comportamientos sociales que conlleva, entonces lo racional sería optar por esto último. Nos hallamos una vez más frente a un conflicto de compromisos: nuestro compromiso con una racionalidad práctica y nuestro compromiso con una racionalidad teórica. No parece que vayamos a ser capaces de averiguar en breve si es más básico

⁵⁰⁹ Strawson 1962, en Fischer & Ravizza 1993: 55-56.

nuestro compromiso con una racionalidad práctica o con una racionalidad teórica, y probablemente tampoco sea necesario averiguarlo.⁵¹⁰ Pero sí se puede pensar que, aunque ambos compromisos son igualmente importantes, en nuestra situación actual hay más razones para mantener uno y no el otro. Pues bien, en este sentido hay una buena razón para mantener nuestro compromiso con la racionalidad práctica que paso a exponer a continuación.

No contamos con leyes científicas manejables, esto es, las estructuras sociales no nos permiten hoy en día actuar como si el determinismo hubiese de suprimir nuestro actual concepto de responsabilidad moral y hacernos prescindir de nuestras emociones reactivas. No hay ninguna incoherencia en la postura compatibilista hoy por hoy: mientras nuestro mundo sea el que es, en que contamos con la atribución de responsabilidad como eje de nuestra vida moral, y dicho mundo funcione correctamente, entonces no es irracional, en ninguno de los dos sentidos de “racional” que se ha distinguido, mantener nuestras actitudes reactivas; pero sí sería irracional dejar de hacerlo al menos desde el punto de vista de la racionalidad práctica. El problema de qué hacer o qué no hacer con nuestra vida práctica no puede anticiparse metafísicamente a su posible vigencia práctica, esto es: la sospecha de que podamos estar siendo incoherentes teóricamente al comportarnos como si fuésemos seres responsables, no es un argumento para que abandonemos lo que he llamado “el mundo participante”. Ceder a este argumento sería claramente irracional, tanto que es manifiestamente complicado hacerlo. Pero no por ello sería, como pretende Strawson, antinatural, pues podrían darse ciertas circunstancias sociales y culturales que ya no lo harían irracional. Deberíamos crear un mundo nuevo para poder abandonar nuestro compromiso con las actitudes reactivas y por ello dicho compromiso parece esencial, irrenunciable. Sin embargo no es así: es lógicamente posible romperlo. Pero es necesario otro

⁵¹⁰ Aunque en general a lo largo de este trabajo he apostado por la prioridad del compromiso práctico, y en Hernández 2003 abogo claramente por éste.

mundo al que ir para poder abandonar el nuestro, es necesario un paradigma nuevo para que pueda haber una revolución científica, y no sólo eso: debe haber cambiado la percepción y la interpretación del mundo, debe haber ciertas evidencias y además el nuevo paradigma debe ser “manejable”, útil.

Ayer, que simpatiza con Strawson, opina que la razón por la cual no querríamos abandonar nuestro mundo participante es porque no poseemos, hoy por hoy, una teoría científica que nos describa suficientemente bien cómo funciona, de manera determinista, la acción humana. Otra forma de decir lo mismo es que normalmente consideramos que cuando algo está determinado, entonces, si poseemos datos suficientes, es predecible. Pero hoy por hoy somos capaces de predecir el comportamiento humano de forma vaga e imprecisa, pero no científica, como bien se pudo ver en *I.3.1.* al confirmarse la hipótesis 3) de la presentación. Reiteradamente los seres humanos nos hemos equivocado cuando hemos intentado predecir el comportamiento ajeno (e incluso a veces el nuestro propio) mediante leyes científicas. Una y otra vez nuestros intentos por adaptar el curso de la historia a leyes ha fracasado, y a menudo de forma sangrienta. Esto no quiere decir que la idea de Ayer de que algún día la ciencia será capaz de llevar a cabo tamaña hazaña sea descabellada (aunque el fracaso del conductismo y doctrinas similares no juega a su favor), pero hay que preguntarse si es racional hoy en día, a cuenta de esa sospecha sobre una posibilidad futura, renunciar a nuestro mundo participante. Y es aquí donde entra en juego con fuerza el argumento pragmatista de Strawson: no parece racional hacerlo.

En definitiva: dejar de lado las actitudes reactivas no sería en absoluto racional porque no contamos con nada con que rellenar el hueco que éstas dejarían, no tenemos ninguna herramienta que nos permita vivir en un mundo diferente del participante. Las cábalas metafísicas del incompatibilista no son esa herramienta, son meras sospechas. Desde una racionalidad práctica no se puede por menos que concluir que lo racional es seguir viviendo en nuestro

mundo participante: funcionando a base de emociones reactivas y disculpas dentro de la actitud reactiva, y a base de veredictos, excusas y exenciones en la actitud reflexiva, ambas partes integrantes de este nuestro mundo participante, el mundo de la responsabilidad moral.

Recapitulación

Según Paul Russell hay dos ideas que son esencialmente constitutivas del punto de vista del incompatibilista:

-“que si la tesis del determinismo es cierta, entonces tenemos razones para rechazar y repudiar las actitudes y prácticas (establecidas) asociadas a la responsabilidad moral sobre la base de que son incoherentes y están injustificadas”⁵¹¹

-“que si tenemos razones para suspender o abandonar las actitudes y prácticas asociadas a la responsabilidad moral, entonces somos, psicológicamente o prácticamente hablando, capaces de hacer tal cosa”⁵¹²

Como se ha visto, el naturalismo se centra sobre todo en negar la segunda instancia y precisamente por ello fracasa, porque no hay modo de averiguar aquello de lo que somos o no capaces los seres humanos psicológicamente hablando. La razón de esta imposibilidad es que el ser humano está muy lejos de ser un animal puramente natural, lo cual lo hace suficientemente maleable como para que sea muy difícil poner límite a sus capacidades. En cualquier caso el naturalismo, de invalidar también la primera instancia, la haría tan sólo de forma accesorio. Ahora, la ventaja de la perspectiva de Strawson es que deja de lado el affaire metafísico y se fija en lo práctico, en los hechos. Sin embargo, de las dos bases del incompatibilismo, la que constituye realmente una tesis metafísica es la que se empeña en falsar Strawson mediante su estrategia naturalista. La estrategia racionalista, en cambio, se centra en la primera demanda del incompatibilista en que la

⁵¹¹ Russell, P. 1992: 289.

discusión que se abre es de carácter práctico y dentro de esta discusión, como se ha mostrado en esta parte del trabajo, una definición pragmatista racional del compatibilismo resulta victoriosa. Lo único necesario para rechazar el incompatibilismo es falsar esa primera demanda, y esto lo hace de sobra la estrategia racionalista que muestra: por un lado en qué consiste nuestro funcionamiento práctico coherente (mediante un sistema de excusas y exenciones y actitudes reactivas que habrá ocasión de investigar con más calma en la parte siguiente); y por otro que esa forma que tenemos de funcionar no es menos sino más racional que la alternativa que propone el incompatibilista, pues debería medirze la racionalidad práctica en términos de ganancias y pérdidas para nuestra vida, lo cual hace que la elección compatibilista sea desde un punto de vista práctico, más racional. Basta con esto para rechazar el incompatibilismo de forma definitiva. En la parte anterior este rechazo se basó en consideraciones metafísicas, aquí se ha procedido a encontrar razones morales para confirmarlo. Tal vez el argumento fenomenológico de la primera parte del trabajo no le resulte convincente a algunos filósofos, y a otros les ocurrirá lo mismo con el argumento pragmatista del capítulo anterior. Quiero creer que la suma de los dos es suficientemente persuasiva, y por ello consideraré zanjada ya de forma definitiva la cuestión de la presunta incompatibilidad entre el determinismo por un lado, y la libertad de la voluntad y la racionalidad de la responsabilidad moral por otro.

Resumiendo, en esta parte se ha visto que llevamos adelante una vida participante dentro de la cual, según las circunstancias, dos actitudes son posibles: la reactiva y la reflexiva. La actitud reactiva incluye ciertas actitudes o emociones reactivas, la actitud reflexiva consiste precisamente en la suspensión excepcional de cualesquiera de estas emociones en razón de algo que consideramos excusa o exime de responsabilidad al agente, o porque las

⁵¹² Ibidem.

circunstancias en que es preciso que juzguemos la responsabilidad de dicho agente (véase en un juzgado) preveen que estén ausentes de nuestro juicio los sentimientos. Estas son las bases de nuestro mundo participante y por lo tanto de nuestro funcionamiento como seres responsables dentro de una comunidad moral. Unas bases que coinciden con la hipótesis 4) de este trabajo: “Tenemos expectativas y emitimos juicios de responsabilidad sobre las acciones ajenas y propias basándonos en dichas expectativas”. Y con esto ya se posee la mitad de una teoría de la responsabilidad moral. La otra mitad será desarrollada en lo que resta de trabajo.

III- EL MUNDO PARTICIPANTE

Al final de la parte II he llegado a la conclusión de que el modo en que nos comportamos normalmente en sociedad, haciendo atribuciones de responsabilidad mediante emociones reactivas, forma parte de nuestra materialización de lo que de la mano de Wittgenstein he llamado una “imagen del mundo”. En este apartado voy a tratar de desarrollar los componentes de esa imagen. Hasta aquí he estado hablando de forma un tanto vaga e imprecisa de emociones, actitudes, exenciones y excusas, tal y como lo hace Strawson en su artículo. En lo que sigue voy a intentar elaborar una teoría de la responsabilidad que dé cuenta de forma pormenorizada de todos estos elementos, de todo lo que compone nuestro mundo participante. Será este análisis el que abrirá la vía de una investigación más allá de las conclusiones de Strawson y a la que dedicaré el resto de esta parte ocupándome, primero, de corregir los errores filosóficos más importantes cometidos por Strawson, y de elaborar, a continuación, una versión propia del juego de la responsabilidad moral siguiendo siempre, eso sí, el patrón conceptual del artículo de Strawson. Aspiro, de este modo, a lograr un bosquejo suficientemente completo del hecho de responsabilidad moral incluyendo su carácter emotivo pero también su alcance normativo.

Para ello me ocuparé en III-1. de ahondar en la intuición strawsoniana de que las emociones reactivas son juicios implícitos sobre expectativas. Así, investigaré qué es lo que las separa de la actitud objetiva, a fin de saber cuál es su diferencia específica, y me preguntaré hasta qué punto pueden algunas de ellas ser consideradas como sentimientos morales.⁵¹³

⁵¹³ He de hacer aquí dos advertencias. La primera de ellas es que tan sólo voy a ocuparme de las emociones reactivas, así pues dejo de lado el importante debate que ha existido en torno a la teoría de las emociones. Creo que lo que yo voy a desarrollar aquí no haría sino completar alguna teoría de las emociones, es un apéndice, y no una teoría en sí misma. Quien esté interesado en algo más que la emociones reactivas hará bien en consultar, al menos, las siguientes obras: Damasio,

En III-2. continuaré con el tema de las emociones reactivas, pero me centraré menos en su definición y más en su funcionamiento en nuestra vida diaria. Trataré de hacer una taxonomía de las mismas y de dar una respuesta definitiva a la cuestión de su alcance moral.

A continuación, en III-3., atenderé a las bases culturales y lingüísticas que dotan de significado a nuestras atribuciones de responsabilidad. Insistiré entonces en la base cultural de nuestras emociones reactivas y analizaré de forma pormenorizada cuáles son las principales categorías en que se ordenan lo que solemos admitir como excusas o exenciones.

Para terminar, en III-4. retomaré la cuestión del significado cognitivo de nuestras emociones reactivas como juicios sobre expectativas, pero en esta ocasión desarrollaré con más atención lo que estos juicios tienen de expresión de la existencia de una comunidad moral de la cual todos somos miembros y dentro de la cual la responsabilidad moral cobra todo su sentido.

Por fin, haré balance de todo aquello que, espero, habré conseguido ir desvelando, corregiré ciertas imprecisiones a la luz de esos nuevos avances y con ello daré por terminada mi interpretación fenomenológico-pragmatista de la responsabilidad moral.

Antonio, 1996, *Descarte's Error. Emotion, Reason and the Human Brain*, New York: G. P. Putnam's Sons; Elster, Jon, 1999, *Alchemisties of the Mind. Rationality and the Emotions*, Nueva York: W. W. Norton & Co. Inc.; Gibbard 1990; Nussbaum, Martha C., 1994, *The Therapy of Desire. Theory and Parctice in Hellenistic Ethics*, Princeton: Princeton University Press; Williams, Bernard, 1993, *Shame and Necessity*, Berkeley: University of California. También recomiendo Hansberg, Olbeth, 1996, *La diversidad de las emociones*, México: Fondo de Cultura Económica, en él hay un capítulo dedicado al resentimiento. En Broncano 2001 el autor ofrece un breve recorrido por las principales teorías de las emociones. La segunda advertencia compete a una emoción concreta: la compasión. Aunque en la última parte de este trabajo hablaré mucho de ella (como empatía), no le voy a asignar un papel central en el entramado de las emociones reactivas, de hecho no la tengo por miembro de las mismas. Ciertamente parece extraño que hable de sentimientos morales y no tenga en cuenta la compasión, que algunos autores consideran la raíz misma de la moral (véase Rousseau), para curar esa extrañeza me remito al principio de esta nota: tan sólo me voy a ocupar de las emociones reactivas, y la compasión no es una de ellas. No dudo que exista un vínculo muy estrecho entre la compasión y la moral, pero no es el mismo vínculo que el que pueda haber entre la moral y una emoción reactiva, y ese es el que me interesa aquí. Quien esté interesado en la compasión y quede consternado por su ausencia aquí podrá saciar sus ansias de empatía con Arteta 1996.

III-1. Ju(z)gar con los sentimientos

“Ángel lleno de alegría, ¿conoces la angustia, / La vergüenza, el remordimiento, los sollozos, los problemas (...)?”⁵¹⁴

Son muchas las publicaciones que han recogido las ideas de Strawson, no tantas aquellas que han desarrollado de forma pormenorizada teorías acerca de la responsabilidad moral llevando las intuiciones de Strawson más allá. Así pues, se tiende a considerar que el artículo de Strawson da cuenta de los aspectos esenciales de un normativismo de corte naturalista. No obstante ciertos autores, aspirando a encontrar una teoría general satisfactoria acerca del concepto de responsabilidad moral, han pensado que era necesario sofisticar las ideas de Strawson y corregir ciertos errores, y esta tarea es la que me va a ocupar también en este apartado.

Sin duda la aportación más importante en este sentido la constituye el *Responsibility and the Moral Sentiments* de Robert Jay Wallace⁵¹⁵, de quien se puede decir que, si bien le da un enfoque cognitivista casi ausente en el artículo “Freedom and Resentment”, es el continuador más fiel de la perspectiva de Strawson. En su obra hace ciertas consideraciones y ciertas críticas que no puedo dejar de tener en cuenta aquí. El objetivo principal de Wallace es construir una teoría general de la responsabilidad moral de corte strawsoniano, pero más detallada y precisa que la del propio Strawson, y por ello se dedica en su obra a atar ciertos cabos que Strawson deja sueltos y a convertir sus intuiciones en una construcción filosófica más sistemática y coherente.

Las principales fallas de la argumentación strawsoniana, que Wallace intenta subsanar, son desveladas por Jonathan Bennett en su influyente

⁵¹⁴ Baudelaire, “Reversibilidad”, en Baudelaire 1991: 208. He preferido hacer mi propia traducción

⁵¹⁵ Wallace 1998.

artículo “Accountability”⁵¹⁶. Muchas son las cosas de las que se ocupa Bennett en su artículo, que básicamente constituye un intento por poner en relieve la importancia de las ideas de Strawson para el debate del determinismo y la libertad, pero entre todas ellas aquí interesa esencialmente examinar su crítica a dos aspectos de la teoría de Strawson que constituyen serias deficiencias dentro de su ingeniosa solución al problema libertad/determinismo. Según Bennett estos dos aspectos, íntimamente ligados, en los que falla la teoría de Strawson son los siguientes:

-Strawson define demasiado borrosamente las actitudes reactivas, no consigue dar con un criterio que permita decidir claramente qué se debe y qué no se debe considerar una actitud reactiva.

-Strawson defiende que sólo las actitudes reactivas que llama “análogas vicarias” son morales, sin embargo da la impresión de que, en tanto todas las actitudes reactivas son fruto de nuestra actividad como agentes responsables, se podría considerar que todas ellas pertenecen igualmente al terreno de la moral.

Estos problemas que señala Bennett son sin duda los más interesantes de analizar a la hora de construir una teoría strawsoniana de la responsabilidad moral más potente, no obstante se puede considerar que existe otro error en la argumentación strawsoniana que, si bien es más marginal, no deja de ser molesto porque tiene la apariencia de una contradicción.⁵¹⁷ Así pues, antes de lanzarme a sofisticar su teoría de la mano de Bennett y Wallace, intentaré superar ese escollo de la argumentación de Strawson. Una vez hecho esto me centraré en la búsqueda de un criterio que permita distinguir las actitudes

⁵¹⁶ Bennett 1980, en Van Straaten 1980.

⁵¹⁷ Véase, una vez más, la nota 422 y el texto al que corresponde.

reactivas, y a continuación me ocuparé de aclarar, en la medida de lo posible, los borrosos límites de la moral.

III.1.1. La actitud objetiva como posibilidad teórica

Cuando en su artículo Strawson habla de la actitud objetiva señala un hecho que parece claramente incompatible con su idea de que dicha actitud sólo puede darse en aquellos casos en que vemos al otro como un “incapacitado” desde el punto de vista moral: Strawson dice que hay momentos en que nos es posible adoptar la actitud objetiva para con agentes plenamente responsables. En efecto, Strawson sostiene que “las actitudes reactivas personales en general, tienden a ceder su sitio, y, consideran las personas civilizadas, deberían ceder su sitio, a las actitudes objetivas, siempre y cuando el agente sea visto como excluido de las relaciones humanas adultas ordinarias por una anormalidad psicológica profundamente arraigada —o simplemente por ser un niño. Pero no puede ser consecuencia de ninguna tesis que no sea ella misma contradictoria que la anormalidad es la condición universal.”⁵¹⁸ Sospechosamente, al mismo tiempo que Strawson afirma esto sostiene también que

“la actitud objetiva no es solamente algo en lo que tendemos a caer en casos como esos, en que las actitudes participativas se ven parcial o totalmente inhibidas por anormalidades o por inmadurez. Es también un recurso que está disponible en otros casos. Miramos con un ojo objetivo el comportamiento compulsivo del neurótico o el molesto comportamiento de un niño pequeño, pensando en términos de tratamiento o educación. Pero *podemos* a veces mirar con algo parecido a ese ojo el comportamiento de la persona normal y madura. *Tenemos* ese recurso y podemos usarlo a veces: como un refugio de las tensiones de la implicación; o como una ayuda para la política; o simplemente por curiosidad intelectual.”⁵¹⁹

La explicación que da Bennett de esta aparente contradicción consiste en señalar que esta posibilidad de adoptar la actitud objetiva con respecto a

⁵¹⁸ Strawson, P. F. 1962, en Fischer & Ravizza 1993: 53-54.

⁵¹⁹ Strawson, P. F. 1962, en Fischer & Ravizza 1993: 52.

alguien normal está “siempre abierta teóricamente”⁵²⁰, esto es, que no es siempre sencillamente incorrecto –aunque a veces sea psicológicamente imposible– disipar nuestros sentimientos reactivos refugiándonos en la objetividad. Podemos considerar a alguien como ‘un caso’ sin creer o pretender que está mentalmente enfermo, etc.”⁵²¹ Según Bennett, la teoría de Strawson permite discriminar en qué ocasiones “*no sería incorrecto* culpar a una persona”⁵²² pero no nos dice con ello que en esas ocasiones “*sería incorrecto no* culparla”.⁵²³ La teoría de Strawson se limita a mostrar que hay ocasiones en que no resulta apropiado sentir actitudes reactivas (bajo ciertas condiciones que eximen o excusan de responsabilidad), pero en ningún caso Strawson habla de circunstancias que obliguen a sentir las actitudes reactivas. En cierto sentido cabe decir que, según Strawson, las actitudes reactivas son hasta cierto punto “libres”, no están directamente adscritas a actitudes proposicionales concretas, aunque hay cierto tipo de actitudes proposicionales que hacen desaparecer esos sentimientos, que parecen ser incompatibles con ellos. Por ello nuestros sentimientos reactivos son maleables y se nos abre la posibilidad de, en ciertas ocasiones, dejarlos de lado y adoptar de forma teórica una actitud objetiva para con alguien que normalmente nos impulsaría a tener esos sentimientos, pero al que en algunos casos “como un refugio de las tensiones de la implicación; o como una ayuda para la política; o simplemente por curiosidad intelectual”⁵²⁴ juzgamos que nos conviene considerar con un “ojo objetivo”.

⁵²⁰ En su artículo Bennett explica que usa “teóricamente” como lo contrario de “con una finalidad práctica” entendiendo esto último como “esperando obtener cierto resultado práctico”. Más adelante, en *III.1.2.*, retomaré este punto.

⁵²¹ Bennett 1980, en Van Straaten 1980: 24.

⁵²² *Ibidem.*

⁵²³ *Ibidem.*

⁵²⁴ Strawson: 1962, en Fischer & Ravizza 1993: 52.

Unas distinciones que realiza Lawrence Stern en su artículo “Freedom, Blame, and Moral Community”⁵²⁵ permitirán sin duda comprender mejor este aspecto. Stern señala que la actitud objetiva posee dos elementos que es necesario distinguir, a saber: en la actitud objetiva hay por un lado una ausencia de sentimiento y por otro lado un esfuerzo “por obtener resultados de la persona con la que estamos tratando”.⁵²⁶ Pues bien, Stern señala que, aunque Strawson tiene razón al insistir en la presencia de estos dos aspectos en la actitud objetiva, es posible sentir los diferentes sentimientos implicados en nuestras relaciones interpersonales corrientes y a pesar de ello “estar atento a un resultado”.⁵²⁷ Es cierto que lo normal es que una persona que esté sobre todo pendiente de obtener cierto resultado de otra tienda a apartar sus sentimientos, pero lo hará no porque ello sea inevitable, porque sea condición necesaria para buscar ciertos resultados el prescindir de los sentimientos, sino por considerar sencillamente inconveniente la obediencia a los sentimientos, de escasa ayuda: es más sencillo manipular a alguien si no dejamos que los sentimientos influyan en nuestra visión de aquel a quien pretendemos manipular.

La conclusión que de esto extrae Stern es que la ausencia de sentimientos es, en realidad, un aspecto vicario de la actitud objetiva, “lo que es esencial es la concentración en los resultados”.⁵²⁸ En este sentido se puede comprender la idea de Strawson de que a veces es posible adoptar la actitud objetiva para con un agente plenamente responsable, aunque sin embargo no es posible hacer de ella nuestra situación normal: nos es muy difícil dejar de lado nuestros sentimientos por cortos espacios de tiempo y parece prácticamente imposible, diría Strawson, hacerlo de forma continuada; ahora, no por ello somos menos capaces de, en algunas ocasiones, comportarnos

⁵²⁵ Stern: 1994.

⁵²⁶ Ibidem: 74.

⁵²⁷ Ibidem.

para con cualquier persona adoptando la actitud objetiva en vistas a obtener cierto resultado de ella.

En fin, el hecho de que sea imposible dejar de lado de forma continua nuestros sentimientos no implica que no podamos comportarnos de forma manipuladora para con cualquier persona y por ello es posible aquello que dice Strawson: la actitud objetiva no siempre está reservada para el enfermo o el niño, “*podemos* a veces mirar con algo parecido a ese ojo el comportamiento de la persona normal y madura”.⁵²⁹

III.1.2. Las emociones reactivas

Es necesario fijarse ahora en uno de los principales problemas del artículo de Strawson: la excesiva borrosidad con que su autor define aquello en que consiste una “actitud reactiva”, una borrosidad que parece asentarse en el prejuicio típicamente filosófico de pensar que, si algo tiene que ver con los sentimientos, entonces es prácticamente inefable. Tal vez no esté de más hacer un paralelismo entre la incapacidad de Strawson para hacer una clasificación algo ordenada de las actitudes reactivas y el error que comete Wittgenstein en las *Investigaciones Filosóficas* al subrayar que hay “*innumerables géneros de empleo de todo lo que llamamos ‘signos’, ‘palabras’, ‘oraciones’*”.⁵³⁰ Así, seguramente se le podría decir a Strawson algo parecido a lo que comenta Searle a propósito de Wittgenstein: “*Más bien la ilusión de los usos ilimitados del lenguaje se origina por una falta de claridad enorme sobre lo que constituyen los criterios para delimitar un juego o un uso del lenguaje de otro.*”⁵³¹ En efecto Strawson, al igual que Wittgenstein, no se preocupa de buscar ningún criterio clasificador y se conforma con citar algunos casos de

⁵²⁸ Ibidem.

⁵²⁹ Strawson 1962, en Fischer & Ravizza 1993: 52.

⁵³⁰ Wittgenstein 1988: 39.

actitudes reactivas, lo cual le conduce a considerar que, una vez más por un prejuicio filosófico muy arraigado, puesto que se le ocurren fácilmente cuatro o cinco casos, entonces en total debe haber innumerables casos más. Strawson no precisa cuales serían las características que harían de una determinada actitud o de un determinado sentimiento una actitud reactiva, simplemente hace una enumeración de lo que para él serían las actitudes reactivas más representativas, a saber: “cosas tales como la gratitud, el resentimiento, el perdón, el amor y los sentimientos heridos”.

La actitud objetiva y la investigación teleológica

Johnathan Bennett ha sido el primero en desvelar esta carencia de la teoría strawsoniana y en intentar solucionarla, a pesar de que él mismo no considere que la borrosidad con que Strawson define las actitudes reactivas sea un problema grave. Para él la explicación strawsoniana “podría cumplir su objetivo principal sin usar ‘reactivo’ o cualquier sustituto para ello, y meramente hablando por separado de ‘resentimiento’, ‘gratitud’, ‘indignación’ y otras pocas [actitudes más]”.⁵³² Esto es así porque la intención de Bennett en su escrito es sobre todo insistir en la importancia que tiene el artículo de Strawson por acercar el problema de la responsabilidad moral al terreno de los sentimientos y poner así fin a una “sobreintelectualización” de la cuestión. Ahora, a pesar de que para el propio Bennett la falta de un criterio de distinción para las actitudes reactivas no resulte ser un problema, aquel no deja de tener presente que sí podría representarlo para cualquier otro filósofo y por ello no renuncia a buscar una definición precisa de aquello a lo que puede llamarse una “actitud reactiva”.

Su intento le lleva a la siguiente fórmula: “La actitud de x hacia y es reactiva *si y sólo si* es una actitud favorable o desfavorable que no podría

⁵³¹ Searle, John “Una taxonomía de los actos ilocucionarios”, en Valdés 1995: 476.

⁵³² Bennett 1980, en Van Straaten 1980: 39.

explicar el que x emprendiese una investigación teleológica sobre cómo funciona y .”⁵³³ Para comprender esta definición hay que atender a qué entiende Bennett por “investigación teleológica” y esto debe hacerse retomando la discusión de más arriba acerca de la actitud objetiva. Se dijo que Bennett defiende que dicha actitud es posible en cualquier ocasión y con respecto a cualquier persona aunque sólo de forma “teórica”. Precisamente, en aquellos casos en que la actitud objetiva sólo puede ser adoptada de forma teórica, entonces no es posible enfrascarse en una investigación teleológica, esto es, “una investigación que aspira a realizar cierto fin práctico, y no a la mera adquisición de conocimiento y comprensión”.⁵³⁴ Se comprenderá mejor este punto a través de un ejemplo. El caso más claro de investigación teleológica, que cita Strawson en su artículo como característico de la actitud objetiva, sería el de un psicoterapeuta con su paciente. En efecto, el terapeuta se interesa por el comportamiento del paciente desde un punto de vista neutral, sin emitir juicios morales y por lo tanto sin sentir ningún tipo de actitud reactiva. El objetivo del terapeuta es descubrir una serie causal de acontecimientos que dé cuenta del comportamiento patológico del paciente con la única idea de realizar en consecuencia un tratamiento que cambie dicho comportamiento. No deja de ser cierto que un médico en presencia de algún paciente puede llegar a verse invadido por ciertos sentimientos que tal vez no dudaríamos en calificar de reactivos, sin embargo los obviará porque su interés como médico en el paciente no es el de “participar con él en una relación interpersonal ordinaria”, sino el de conseguir manipularlo hasta cambiar su carácter, o cuando menos lograr explicar científicamente su conducta. “La investigación teleológica busca verdades que puedan ser aplicadas tanto en la

⁵³³ Ibidem: 38.

⁵³⁴ Ibidem: 36.

persecución de fines que son vistos como inherentemente deseables como en vistas a algún bien inherente subsecuente.”⁵³⁵

En definitiva, parece que la investigación teleológica conlleva las dos cosas que Stern señala como propias de la actitud objetiva: por un lado una ausencia de sentimientos (útil a la hora de adquirir una explicación científica u objetiva de una acción) y por otro lado un intento por obtener algún tipo de beneficio de cierta persona. Así pues, teniendo en cuenta lo visto en el apartado anterior, la investigación teleológica no parece ser otra cosa que el comportamiento tipo que conlleva la actitud objetiva. Por ello Bennett, con su caracterización de las actitudes reactivas basada en la imposibilidad de darse la investigación teleológica, lo único que estaría diciendo es que dichas actitudes no son sino lo contrario de la actitud objetiva. Y así lo admite él mismo: “mi propuesta de condición necesaria y suficiente para que una actitud sea ‘reactiva’ no es ni útil ni satisfactoria. Sólo nos dice que la reactividad es incompatible con cierto poder causal o explicativo (...). Esto no nos dice lo que es la reactividad, sino solamente nombra una de sus propiedades (...).”⁵³⁶ En definitiva: Bennett suministra más bien un criterio para distinguir entre *la* actitud reactiva y *la* actitud objetiva, y no un criterio para reconocer *una* actitud reactiva (una atribución emotiva de responsabilidad).⁵³⁷

Sobre las expectativas

Este infructuoso esfuerzo de Bennett a la hora de definir las actitudes reactivas no deja de resultar extraño si se tiene en cuenta que sí hay algo

⁵³⁵ Ibidem: 37.

⁵³⁶ Ibidem: 38.

⁵³⁷ Ahora, aunque este no era el objetivo perseguido, tampoco está exento de importancia el haber conseguido dar con un criterio que permita distinguir claramente en qué consiste la actitud objetiva. De hecho ahora se comprende mejor aquello que, como se vio más arriba, afirma Stern sobre el carácter secundario de la ausencia de sentimientos dentro de la actitud objetiva: si en la actitud objetiva suelen estar ausentes los sentimientos es porque ésta se basa en un análisis objetivo, casi científico, del agente, análisis para el cual los sentimientos no importan, y no porque

parecido a un criterio para distinguirlas en el artículo de Strawson, a saber: una actitud reactiva se da basándose en cierta expectativa o demanda de buena voluntad. Ahora, sin dejar de ser cierto que esa demanda o expectativa puede constituir un buen criterio, no es menos cierto que Strawson no hace demasiados esfuerzos a la hora de explicar más precisamente en qué podría consistir esa demanda o expectativa, y tal vez por ello mismo Bennett prescinde de ese criterio.

Quien no prescinde de ello y desarrolla esta idea hasta hacerla funcional es Wallace en su *Responsibility and the Moral Sentiments*. Según él “hay una conexión esencial entre las actitudes reactivas y una forma precisa de evaluación, o cuasi evaluación, a la que referirse como tener cierta expectativa de (o hacer una demanda a) una persona. Esta forma de evaluación no es anterior a las actitudes reactivas, sino que más bien se define en sus términos: tener cierta expectativa de alguien, mantengo, es ser susceptible a las actitudes reactivas en nuestras relaciones con esa persona.”⁵³⁸ En principio esta idea es tan sólo una guía, pues con ella Wallace todavía no está diciendo nada nuevo, aún es necesario hacer ciertas distinciones. Lo primero que señala Wallace es hay que distinguir entre dos formas de entender lo que es una expectativa.

Una primera forma de entender una expectativa no está vinculada en absoluto a la moral o a las emociones reactivas.⁵³⁹ Estar expectante con respecto a algo puede ser esperar que ese algo ocurrirá porque pensamos que hay muchas probabilidades de que vaya a ocurrir. Este tipo de expectativa no está esencialmente vinculada a respuestas emocionales. Por ejemplo, antes de empezar el curso la profesora que decida explicar los *Prolegómenos* de Kant a sus alumnos pensará que es muy probable, esto es, *esperará* de ellos, que no sean

precisamente la actitud objetiva consista en la falta misma de sentimientos. La actitud objetiva sería pues la posibilidad de realizar una investigación teleológica del agente.

⁵³⁸ Wallace 1998: 19.

⁵³⁹ A partir de aquí hablaré de “emociones reactivas” y de “actitudes reactivas” indistintamente, tal y como lo hace Wallace.

capaces de comprender esa obra por sí solos. Independientemente de esta expectativa es posible que la profesora empiece el curso con ilusión, o con miedo o con cualesquiera otros sentimientos.⁵⁴⁰ Según esta forma de entender la expectativa, ligada a un cálculo implícito de probabilidades, es posible, aunque no indispensable, que sintamos ciertas emociones, pero no puede considerarse que se trate de emociones reactivas.

Pero existe otra forma de entender lo que es tener una expectativa, y bajo esta otra forma las expectativas aparecen como esencialmente ligadas a cierto tipo de emociones que no se podría dudar en considerar reactivas. Así, siguiendo con el ejemplo anterior, la profesora *esperará* en este otro sentido que los alumnos no mentirán, o no harán trampa, o no se comportarán en clase como auténticos energúmenos. Este es el sentido de la expectativa que interesa aquí. Cuando se habla de expectativas en relación con las emociones reactivas no se está hablando de esperar algo de algo o alguien, ligando la expectativa a la idea de probabilidad. En efecto, sin duda la profesora piensa que es muy probable que sus alumnos no sean unos mentirosos, pero esta referencia a la probabilidad no da cuenta de lo que es esencial a este tipo de expectativa que la profesora tiene, y que es muy diferente de aquella otra según la cual espera cabalmente que sus alumnos necesiten de su ayuda para comprender los *Prolegómenos*. Según Wallace, en lo que respecta a esperar algo de alguien en el sentido que aquí interesa, “el elemento crucial, sugeriría, es actitudinal: tener cierta expectativa de alguien es esencialmente ser susceptible a cierto rango de emociones en el caso de que la expectativa no se vea colmada, o creer que la violación de la expectativa haría apropiado que uno fuese sujeto de esas emociones.”⁵⁴¹ A esto decide llamarlo Wallace la *postura cuasi-evaluativa*, y es lo que funciona como criterio que define las actitudes reactivas, que no serían

⁵⁴⁰ Este ejemplo es el que utiliza el propio Wallace.

⁵⁴¹ Wallace 1998: 21.

según Wallace sino las emociones propias a esa frustración de una expectativa.⁵⁴²

Así pues, las emociones reactivas están, como ya intuyó Strawson, esencialmente ligadas a demandas o expectativas. Dicho de modo más preciso: cuando sentimos una emoción reactiva es porque “en una situación concreta de acción un *requerimiento* práctico o una *prohibición*”⁵⁴³ no han sido respetados. Así, toda expectativa debería poderse formular lingüísticamente en términos que expresasen un requerimiento o una prohibición: o bien usando expresamente los conceptos de requerimiento o prohibición, o bien bajo la forma de un imperativo. De este modo se puede explicar aquello en que consiste, por ejemplo, el que yo me sienta indignado porque mi novio o novia me ha sido infiel como que estoy indignado porque pienso que mi novio o novia ha violado la siguiente demanda: “No me serás infiel”. Este ejemplo muestra que las emociones reactivas se corresponden con cierta actitud proposicional, en concreto con la “creencia de que cierta demanda ha sido violada”.⁵⁴⁴ Pero cuando sentimos cierta emoción reactiva no hay ninguna prioridad entre ella y la creencia, ambas son la misma cosa: sentir determinada emoción reactiva es estar en cierto estado proposicional dominado por la postura cuasi evaluativa. Aunque a veces no seamos capaces de expresarla lingüísticamente, nuestras emociones reactivas deben ser siempre explicables por una creencia en que se hace patente que se ha roto una expectativa que teníamos puesta en alguien. Así, Wallace ha conseguido dar con un criterio que define las emociones reactivas: “(...) hay una característica que es esencial a las emociones reactivas, y que debe ser adoptada para distinguirlas de otros tipos

⁵⁴² Insisto en que esa es la idea de Wallace, a mí no me parece que las emociones reactivas estén directamente relacionadas con expectativas frustradas, sino también con expectativas cumplidas. Retomaré esta cuestión en *III.1.3.* y en *III.2.3.*

⁵⁴³ Wallace 1998: 22.

⁵⁴⁴ *Ibidem*: 24.

de actitudes. Es su conexión con las expectativas, la conexión que he descrito en términos de la postura cuasi evaluativa de tener expectativas de alguien.”

Y ahora que ya he dado con un criterio, debería ocuparme de aplicarlo y de sacar a la luz la clase de las emociones reactivas.

La clase de las emociones reactivas

Según Wallace, tanto Strawson como Bennett hacen demasiado grande el campo de las emociones reactivas incluyendo cosas tales como el amor recíproco, los sentimientos heridos y muchas otras⁵⁴⁵, y precisamente por ello tienen tantos problemas para encontrar qué es lo que caracteriza a las emociones reactivas. Tal y como ellos tratan la cuestión, casi cualquier emoción que se dé en relaciones interpersonales parece ser reactiva. “Lo que uno encuentra en Strawson y sus comentaristas es una tendencia a equiparar dos distinciones que se ven mejor separadamente: (1) la distinción entre actitudes reactivas y otras actitudes que uno pueda tomar hacia las personas; y (2) la distinción entre (en palabras de Strawson) ‘la actitud (o rango de actitudes) de participación en relaciones humanas, de un lado, y lo que puede llamarse la actitud (o rango de actitudes) objetiva hacia otro ser humano, de otro lado.’”⁵⁴⁶ Así pues, parece que Strawson considera que son lo mismo la actitud participante y la actitud reactiva, así como la actitud objetiva y la actitud reflexiva. No en balde en su intento por definir las actitudes reactivas Bennett acababa suministrando un criterio que las caracteriza precisamente como lo contrario de la actitud objetiva, o sea, como lo mismo que la actitud participante. Pero más arriba, en *II.4.1.*, como se ve claramente en el gráfico de la página 288, ya señalé que, dado que la responsabilidad es un trasfondo, una forma de vida, no puede considerarse que cada vez que uno adopta una actitud objetiva para con alguien, entonces esté ausentándose temporalmente de la

⁵⁴⁵ Así lo hace también Susan Wolf en Wolf 1981.

⁵⁴⁶ Wallace 1998: 27. Más adelante, en *III.2.1.* insistiré en este punto.

participación normal con otros seres humanos. No, el mundo participante contempla la posibilidad, en las ocasiones apropiadas, de juzgar a alguien objetivamente y no sólo reactivamente. Así, lo que permita caracterizar al mundo participante no debería ser lo mismo que permita caracterizar a la actitud reactiva. Wallace es consciente de este punto: formar parte del mundo participante es saber participar de un juego de demandas y expectativas, adoptar la actitud reactiva y sentir las emociones que le son propias es una forma, entre otras, de participar en ese juego, y no la única forma de hacerlo. Por otra parte, una cosa sería *la* actitud reactiva, una disposición, una forma de juzgar; y otra *las* actitudes o emociones reactivas, juicios concretos. Emociones eso sí, que, aun siendo varias, no son cualquier emoción (no cabrían el amor y los sentimientos heridos)⁵⁴⁷, ni tampoco se dan de cualquier manera (sino como expresión de la violación de una demanda). Así, Wallace opina que puede distinguirse entre diferentes tipos de emociones, y que reduciendo la clase de las emociones reactivas aplicando el criterio de la postura cuasi evaluativa se gana mucho en claridad. Aplicando dicho criterio a las emociones que menciona Strawson, según Wallace sólo sobreviven tres: el resentimiento, la indignación y la culpa, “otros candidatos propuestos, como la simpatía, el amor, los sentimientos heridos, y la vergüenza, simplemente no cuentan como actitudes reactivas (o por lo menos como casos no ambiguos de tales actitudes)”⁵⁴⁸.

Aparte de la claridad, el interés de reducir la clase de las emociones reactivas es que así se obtiene una mejor interpretación y una defensa más sólida de dos aspectos en los que insiste Strawson en su artículo, a saber: “que

⁵⁴⁷ Esto no quiere decir que, por ejemplo, entre dos enamorados no haya formas especiales de darse ciertas emociones reactivas: el que dos personas formen una pareja conlleva un nuevo paquete de expectativas que cumplir o que frustrar, y metamorfosea ligeramente algunas emociones reactivas (la envidia, por ejemplo, toma la forma de los celos). Creo que la estructura base sigue siendo la misma, cambian las formas y los compromisos, no creo que ello conlleve una crítica al modelo que estoy tratando de construir, sólo me obliga a hacer la muy necesaria precisión de que ese modelo ha de, y pretende, ser sensible al contexto.

las actitudes reactivas se oponen a la postura objetiva del científico o del terapeuta o del ejecutor de la política social”⁵⁴⁹ y “que las actitudes reactivas son naturales, y los seres humanos son objeto de ellas inevitablemente”.⁵⁵⁰ En efecto, por un lado, al vincular la reactividad al hecho de que hacemos cierta demanda nos estamos separando del papel del terapeuta: cuando el médico atiende a su paciente es precisamente porque juzga que no tiene sentido demandarle nada, la utilidad de la actitud objetiva en ese caso es precisamente lograr de alguien que se desarrolle de tal forma que cobre sentido poner alguna expectativa en él. Como señalé más arriba: “Un terapeuta, por ejemplo, podría consistentemente sentir cierto tipo de amor o simpatía por sus pacientes; pero si llegara a sentir resentimiento o indignación hacia sus pacientes, al hacerlo habría cesado de relacionarse con ellos de la forma objetiva característica de su rol profesional.”⁵⁵¹

Por otro lado, recuérdese que el argumento de Strawson para defender que las actitudes reactivas son naturales consiste en sostener que poseer ciertas actitudes reactivas es lo mismo que involucrarse en relaciones interpersonales y, puesto que estas relaciones interpersonales son inevitables, así también lo son las actitudes reactivas. La nueva visión que Wallace da de las actitudes reactivas es más interesante porque no convierte exactamente a las actitudes reactivas en emociones naturales sin más, sino que las hace encajar con un constructivismo muy razonable ⁵⁵² según el cual, si el actuar mediante expectativas es un rasgo indispensable para el desarrollo moral dentro de ciertas sociedades, entonces ese actuar mediante expectativas acaba conduciendo a la aparición de ciertas emociones reactivas de forma natural.

⁵⁴⁸ Wallace 1998: 30.

⁵⁴⁹ Ibidem.

⁵⁵⁰ Ibidem: 31.

⁵⁵¹ Ibidem: 30.

⁵⁵² En concreto Wallace reproduce la argumentación de John Rawls en *A Theory of Justice*, Chapter VIII, secs. 70-74. Desgraciadamente no puedo repetirla aquí porque me llevaría demasiado tiempo y me separaría del objetivo fundamental del trabajo.

Naturalmente esta es una idea de “natural” no muy naturalista (en el sentido más fuerte que ya critiqué más arriba, no en el de un naturalismo pragmatizado como el que ha sido adoptado en este trabajo). El propio Wallace indica que teniendo en cuenta la postura cuasi evaluativa, que vincula las emociones reactivas a ciertas actitudes proposicionales con forma de expectativa, “tenemos que suponer que [las emociones reactivas] tienen que tener una dimensión cognitiva: en particular, parece que una persona objeto de una actitud reactiva debe tener cierto tipo de creencia evaluativa, una que figure en la explicación de su estado emocional”.⁵⁵³

Así pues, cabe defender que vivir en base a las emociones reactivas es incompatible con la actitud objetiva y que además es, si no directamente inevitable, sí al menos difícilmente concebible no hacerlo, y todo ello sin confundir nuestras emociones reactivas con el conjunto de todas las emociones tal y como hace Strawson. “(...) La interpretación restringida de las actitudes reactivas (...) ofrece una lectura plausible (aunque modificada) de la idea de Strawson de que la gente es naturalmente susceptible a las actitudes reactivas. También clarifica y agudiza la oposición entre la postura objetiva y las emociones reactivas, y evita la confusión conceptual que surge cuando las emociones reactivas se toman como coextensivas de las emociones que uno siente en virtud de su participación en relaciones interpersonales. Sólo por estas razones, parece preferible a la interpretación inclusiva de la reactividad favorecida por Strawson y sus comentaristas.”⁵⁵⁴

En definitiva: las emociones reactivas forman parte del conjunto de emociones que se dan cuando nos involucramos con los demás en relaciones interpersonales; por ello en principio son incompatibles con cualquier tipo de investigación teleológica, y por lo tanto incompatibles con la actitud objetiva. No obstante, no cualquier emoción incompatible con la actitud objetiva es

⁵⁵³ Wallace 1998: 33.

⁵⁵⁴ Ibidem: 32.

reactiva: son emociones reactivas aquellas que llevan consigo la “creencia de que cierta demanda ha sido violada”. Inequívocamente se pueden entonces señalar como emociones reactivas las siguientes: el resentimiento, la indignación y la culpa.

Quedaría así resuelto el primer problema importante con el que se encuentra uno al analizar la teoría de Strawson, a saber, el de la excesiva borrosidad con que éste define las actitudes reactivas. Toca ahora ocuparse del que ha sido señalado como su segundo gran problema, esto es, el de encontrar un criterio para distinguir entre las actitudes reactivas morales y las no morales.

III.1.3. Emociones reactivas morales y no morales

Como ya anuncié más arriba, Bennett considera que uno de los fallos de la teoría de Strawson es que excluye gran parte de las actitudes reactivas del terreno de la moral. Strawson parece cometer el error de pensar que sólo se hacen juicios morales cuando uno mismo no esté implicado en la acción que es objeto de dicho juicio, como si sólo las emociones reactivas que tienen lugar de forma impersonal pudiesen ser consideradas morales por el hecho de ser menos “subjetivas”. En efecto, como bien señala Antonio Valdecantos en su artículo “Emociones responsables”, cuando sentimos indignación se trata de una emoción que “busca ser compartida”.⁵⁵⁵ Ahora, que la indignación lleve consigo “tendencias proselitistas, a menudo muy difíciles de refrenar”⁵⁵⁶, no quiere decir que otras emociones reactivas como la culpa o el resentimiento no puedan llegar a ser compartidas por los demás, aunque en principio, si se sigue a Strawson, haya que considerarlas subjetivas. El hecho de que no hagamos explícito nuestro resentimiento ni nuestro sentimiento de culpa no quiere decir

⁵⁵⁵ Valdecantos 2001: 80.

⁵⁵⁶ Ibidem.

que pensemos que no pueden ser compartidos, es más, generalmente podemos analizar estos sentimientos y decidir si están o no justificados precisamente juzgando que cualquier otra persona en nuestro lugar podría sentir lo mismo. Y esto se parece mucho a lo que es hacer un razonamiento de carácter moral (que, como se ve, está íntimamente ligado con el concebir posibilidades alternativas en la forma en que éstas fueron explicadas en la primera parte del trabajo). La única diferencia existente entre el resentimiento, la indignación y la culpa es de punto de vista, y esa diferencia de punto de vista provoca diferentes manifestaciones sentimentales de una misma atribución negativa de responsabilidad. El resentimiento es una manifestación personal hacia terceros, la indignación es una manifestación impersonal hacia terceros, la culpa es una manifestación personal hacia uno mismo. El que dos de estas manifestaciones sean de facto personales no quiere decir que no sean universalizables: una misma acción me provocará sentimiento de culpa a mí e indignación a un observador externo, ¿habrá que considerar que su juicio es más moral que el mío? Da la impresión de que no. Sólo en el caso de que yo me diese cuenta de que mi sentimiento personal no provoca ningún tipo de sensación equivalente en terceras personas podría yo sospechar que hay algo enfermizo o compulsivo en mi sentir, y entonces sí estaría fuera de la moral, pero en principio no por ser subjetivo deja algo de ser moral porque no por ser subjetivo deja de ser universalizable.⁵⁵⁷

Lo que resulta curioso de la postura de Strawson es que éste tiene una idea muy amplia de aquello a lo que habría que considerar una actitud reactiva, y por lo tanto estaría dispuesto a meter una gran cantidad de actitudes en el saco de la moral. Sin ir más lejos Strawson coincide en señalar el amor recíproco como una actitud reactiva. Esa actitud reactiva sería por definición

⁵⁵⁷ O tal vez debería decir más bien que no por ser un sentimiento particular mío, dicho sentimiento es forzosamente subjetivo. Dejando de lado la moral: cuando digo de una obra de arte

(puesto que es recíproca) o bien personal o bien vicaria, lo cual haría del amor una actitud reactiva moral. Este hecho no puede por menos que interpretarse como un signo de debilidad en los criterios utilizados por Strawson para marcar los límites de la moral, pues resulta bastante contraintuitivo tener que considerar al amor como un sentimiento moral. Esta constatación obliga a no conformarse con los criterios de Strawson para distinguir entre actitudes reactivas morales y no morales. No resulta satisfactorio el criterio establecido por Strawson como representativo de la moral, a saber: una vicariedad en la emoción reactiva que consistiría en su desvinculación con cualquier objeto concreto o en su vinculación a algún tipo de principio general. Así, Bennett explica que según Strawson “una actitud mía es no moral si su base me incluye de forma esencial: me *indigno moralmente* de tu actitud despectiva hacia un benefactor, pero me siento *resentido* por tu actitud despectiva hacia *mi*”.⁵⁵⁸ Pero Bennett considera que existen numerosas emociones reactivas en que uno no desaparece como protagonista y que sin embargo están en el terreno de la moral; y por otra parte hay actitudes reactivas impersonales que sin embargo no son morales (Bennett da el ejemplo de una gratitud irracional que mi amor por la música de Bruckner me haría sentir hacia cualquier amante de su música: ahí claramente se trata de valorar la actitud de alguien hacia una tercera persona, Bruckner, pero no se puede decir que mi valoración sea por ello moral). Por esta razón Bennett propone un criterio para marcar los límites de la moral en las actitudes reactivas que no depende de la posición del sujeto con respecto a la causa de la actitud reactiva, sino más bien de su justificación. Así, distingue entre actitudes “movidas por principios” o “morales” y “no movidas por principios” (aunque no explica suficientemente esta distinción) y llega a la conclusión de que los tres diferentes tipos de actitud reactiva distinguidos por

que es bella y no simplemente que me gusta, estoy llevando a cabo un juicio que dista mucho de ser subjetivo, pues aspira a enunciar una verdad objetiva.

⁵⁵⁸ Bennett 1980, en Van Straaten 1980: 46.

Strawson podrían ser morales, que no hay nada especial en cada uno de ellos que lo haga específicamente moral, y que no por ser “auto-reactiva” (esto es, por versar acerca de uno mismo) es menos moral una actitud reactiva. Así lo admite Strawson en su respuesta a Bennett: “Bennett tiene ciertamente razón al decir que me equivoqué al usar el término “moral” excluyendo de su alcance las actitudes ‘auto-activas’”.⁵⁵⁹ Así pues, la clase de las actitudes reactivas morales resulta ser más extensa aún de lo que suponía Strawson, y se basa en la existencia de principios que justificarían mis actitudes y no en la vicariedad de mis sentimientos. Naturalmente esto no resuelve realmente el problema, pues Bennett da por hecho que todo el mundo entiende en qué consiste una actitud reactiva basada en principios, pero esto no está tan claro. Además, en el apartado anterior se pudo ver que la forma en que Bennett intenta caracterizar las actitudes reactivas no es satisfactoria, lo cual hace doblemente vaga su postura. Por ello una vez más hay que intentar ir más allá e investigar otra posible solución al problema.

Más arriba se pudo ver que Wallace señalaba que, puesto que las emociones reactivas tienen objetos proposicionales (“uno está resentido o indignado *por* algo que alguien ha hecho, o se siente culpable *por* haberle hecho algo a alguien”⁵⁶⁰), entonces deben tener una dimensión cognitiva. Por ello Wallace considera que para distinguir entre emociones reactivas morales y no morales, así como entre las emociones reactivas morales y cualesquiera otras emociones morales, “necesitamos una explicación de la dimensión cognitiva en la actitudes reactivas”.⁵⁶¹ Como ya se vio, los objetos proposicionales que distinguen a las emociones reactivas son creencias acerca de la violación de una expectativa, pues la presencia de tales creencias es lo que explica que se den cualesquiera emociones reactivas. Ahora, esto no quiere decir que tales

⁵⁵⁹ Strawson, P. F., “P. F. Strawson Replies: Reply to Ayer and Bennett”, en Van Straaten 1980: 266.

⁵⁶⁰ Wallace 1998: 33.

creencias sean condición suficiente para que se dé una emoción reactiva concreta: podemos no sentir ningún tipo de resentimiento aunque tengamos la creencia de que cierta expectativa que teníamos ha sido violada; en tal caso pensaríamos que no sería inapropiado sentir resentimiento, pero no que fuera obligatorio.⁵⁶² ¿Qué es lo que hace que, dándose un objeto proposicional del mismo tipo, el agente a veces sienta ciertas emociones reactivas y a veces no, o que con un mismo objeto proposicional un agente sienta ciertas emociones reactivas y otro no? Wallace sugiere que existen diferencias en las posturas de los diferentes agentes con respecto a las expectativas de las que se trate. Un determinado agente adoptará una postura con respecto a sus expectativas, tal que dicha postura justificará el que sienta ciertas emociones reactivas; otro adoptará otra postura, tal que esa en cambio no lo justificará. En este sentido se puede decir que la susceptibilidad a ciertas emociones reactivas (en eso consistiría esa postura que impulsa al agente a sentir o no ciertas emociones reactivas) no explica realmente el que el agente se vea dominado o no por una emoción reactiva, sino que, más bien, esa susceptibilidad, esto es, “el que el agente tiene esa postura se *expresa* en el hecho de que creencias acerca de la violación de ciertas expectativas dan nacimiento a estados de emoción reactiva”.⁵⁶³

La idea de Wallace es que del mismo modo que diferentes posturas con respecto a ciertos objetos proposicionales nos impulsan o no a sentir una emoción reactiva, se podría pensar que son también diferentes posturas con respecto a esos objetos, esto es, diferentes formas de considerarlos, las que hacen que nuestra emoción reactiva tenga o no carácter moral. En este sentido, puesto que toda actitud reactiva se da basándose en ciertas expectativas, según Wallace dependerá del hecho de que esas expectativas

⁵⁶¹ Ibidem.

⁵⁶² Ya traté este punto en III.1.2.

⁵⁶³ Wallace 1998: 34.

tengan o no carácter moral, el que se pueda considerar que una emoción reactiva provocada por ellas sea o no moral: “no hay ninguna razón para suponer que todas las expectativas que tenemos de la gente sean distintivamente de carácter moral, así que deberíamos distinguir entre la creencia de que una expectativa moral ha sido violada y la creencia de que una expectativa no moral ha sido violada”.⁵⁶⁴ De este modo, el que una emoción reactiva sea moral querrá decir que se explica por la creencia de que una expectativa de carácter moral ha sido violada. Ahora, ¿en qué consiste una expectativa de carácter moral?

Según Wallace es muchas expectativas se justifican mediante razones, que en algunos casos estas razones se dan en términos específicamente morales, y que en estos casos se debería considerar que nuestras expectativas son expectativas morales. Wallace decide llamar a estas expectativas morales “obligaciones” y concluye que “deberíamos contar como morales las emociones reactivas cuando están ligadas a obligaciones para las cuales el agente es capaz por sí mismo de dar justificaciones morales; estas justificaciones se basan en razones que explican los esfuerzos del propio agente por obedecer a las obligaciones en cuestión, y suministran términos morales que el agente está preparado para usar a la hora de justificar la misma obediencia por parte de los demás, a los cuales el agente somete a esas obligaciones”.⁵⁶⁵ Así pues, parece que se debería considerar que una emoción reactiva es moral si y sólo si es causada por la violación de una expectativa que seríamos capaces de justificar moralmente.

En conclusión, parece más o menos claro a raíz de esta discusión que el argumento de Strawson para distinguir entre emociones morales y emociones no morales, esto es, el hecho de que el sujeto de las emociones no esté implicado en la acción más que vicariamente, no es un criterio satisfactorio. Si

⁵⁶⁴ Ibidem.

⁵⁶⁵ Wallace 1998: 36.

se considera que, como pretende Bennett, lo que hace que una emoción pueda ser llamada moral tiene que ver con que dicha emoción pueda ser compartida por un observador externo, entonces cualquier emoción reactiva, de estar justificada, sería moral. Sin embargo esto parece ampliar excesivamente el terreno de la moral, y por ello Wallace defiende que tenemos expectativas que pueden considerarse obligaciones y otras que no, y que lo que hace que una emoción reactiva pueda considerarse moral es que esté vinculada a una expectativa que constituya una obligación moral. No parece difícil estar de acuerdo en esto con Wallace, pero desgraciadamente ello se debe a que en realidad lo que está diciendo Wallace desplaza el problema de los límites de la moral, pero no lo resuelve. Una emoción reactiva moral es la que es causada por una expectativa moral, esto es, por una obligación. Una obligación es una expectativa que seríamos capaces de justificar moralmente. ¿Y en qué consiste una expectativa que seríamos capaces de justificar moralmente? Esta cuestión queda sin resolver. Por ello más adelante, en *III.2.3.*, afrontaré de nuevo este problema buscando una solución original a este problema que la principal literatura en torno a las ideas de Strawson creo no ha sido capaz de resolver.

III-2. Anatomía de las emociones responsables⁵⁶⁶

–[E]n la consola vaciló entre marcar un inhibidor talámico (que suprimiría su furia), o un estimulante talámico (que la incrementaría lo suficiente para triunfar en una discusión).

–Si aumentas el volumen de la ira –dijo Iran atenta, con los ojos abiertos– haré lo mismo. Pondré el máximo, y tendremos una pelea que reducirá a la nada todas las discusiones que hemos tenido hasta ahora.

[...] Él suspiró, derrotado por la amenaza.

–Marcaré lo que tengo programado para hoy –examinó su agenda del 3 de enero de 1992: preveía una concienzuda actitud profesional-. Si me atengo al programa –dijo cautelosamente-, ¿harás tú lo mismo? (...).

–Mi programa de hoy incluye una depresión culposa de seis horas –respondió Iran.”⁵⁶⁷

Aunque la reconstrucción que hace Wallace de la teoría strawsoniana es sin duda elaborada y completa, no deja de ser molesto el hecho de que contradiga plenamente ciertas intuiciones de otros seguidores de Strawson que no parecen ser descabelladas. Por un lado Wallace tiene una idea muy restringida de lo que debe ser llamado una “actitud reactiva”, no así Bennett, Susan Wolf o el propio Strawson, y los argumentos que da Wallace para justificar este hecho no resultan plenamente satisfactorios. Por otro lado, con respecto a los criterios para determinar qué emociones reactivas serían morales y cuales no, hay un conato de acuerdo pues la idea de Wallace de que las emociones reactivas morales dependen de obligaciones aparece como una forma sofisticada de aquella idea de Bennett de distinguir entre emociones reactivas movidas y no movidas por principios. No obstante la teoría de Wallace acaba no llegando a ser mucho más clara que la de Bennett, pues no aclara cómo se puede distinguir una obligación.

Es mi idea que estos problemas nacen de ciertas imprecisiones que voy a intentar limar aquí a fin de lograr dar, en la medida de lo posible, con una reelaboración completa y conciliadora de la perspectiva strawsoniana.

⁵⁶⁶ Tomo esta expresión de Valdecantos 2001.

⁵⁶⁷ Dick, *Sueñan los androides con ovejas eléctricas*, 1992: 12.

III.2.1. Actitud participante y sentimientos reactivos

En primer lugar, debo insistir en una distinción señalada por Wallace aunque no aclarada suficientemente. Tanto Strawson como Bennett y Wolf toman como referencia de las actitudes reactivas el hecho de que se den en relaciones interpersonales ordinarias, y esto mismo es lo que les hace acabar considerando que casi cualquier sentimiento es una actitud reactiva. Ahora, siendo fieles a Strawson no se debe olvidar que, *en rigor*, la capacidad para participar normalmente en relaciones interpersonales es estar en la actitud participante. Hay pues preguntarse ahora si estar en la actitud participante es sinónimo de sentir una actitud reactiva, pues esto es algo que no deja claro de forma explícita Strawson.⁵⁶⁸

Si se atiende a Bennett, una actitud es una “postura mental que es *apta para una interacción de cierto tipo*”⁵⁶⁹ y si se considera aquello en que consistiría según esta definición la actitud participante, se diría que *es una postura mental apta para la interacción con otros seres humanos adultos y normales*. O sea, exactamente lo contrario de la actitud objetiva que está dominada por una investigación teleológica. Ahora, el resentimiento, la indignación, la culpa, ¿son realmente actitudes en este sentido? ¿No debería decirse más bien que se trata de sentimientos o emociones que se dan *cuando* estamos en la actitud participante, esto es, con una postura mental apta para la interacción con otros seres adultos y normales? Es más, ¿no se vio más arriba que a veces podemos adoptar la actitud objetiva y sin embargo tener ciertos sentimientos que no dudaríamos en calificar de “reactivos”? Esto es porque lo que llama Strawson “actitudes reactivas” no son realmente actitudes. La explicación a ciertas contradicciones

⁵⁶⁸ Esta cuestión ya fue tratada más arriba, en II.4.1.

⁵⁶⁹ Wallace 1998: 34.

viene de la mano de esta distinción: ⁵⁷⁰ una emoción reactiva es un determinado tipo de emoción que se da puntualmente en aquellas ocasiones en que sentimos frustrada una expectativa. La actitud participante es algo mucho más amplio, es nuestra perspectiva mental normal a la hora de relacionarnos con los demás, es una postura social continua. Así, como tal actitud incluye multitud de clases diferentes de emociones: una de ellas es la de las emociones reactivas, pero no la única.

Así pues, no es que Strawson y Bennett dilaten excesivamente la clase de las actitudes reactivas sino que, como no consiguen dar con un criterio definitorio de “actitud reactiva” acaban confundiendo actitud reactiva y actitud (vida o mundo, tal y como defendí en *III.4.1.*) participante, esto es, la parte con el todo. No es de extrañar que Strawson hable casi siempre de la actitud participante en singular: ⁵⁷¹ no hay varias actitudes participantes y sin embargo sí hay muchas actitudes reactivas. La participación es más bien algo en lo que se está o en lo que se vive; o si se quiere, más en la perspectiva de Strawson, algo que se tiene de modo casi consustancial. Por el contrario las actitudes reactivas, como en realidad son emociones, son puntuales, múltiples y cambiantes; tanto, que incluso pueden aparecer cuando adoptamos una actitud objetiva. ⁵⁷² “(...) [L]a categoría de la reactividad no representa una forma ilustrativa de agrupar emociones, siempre que el contraste entre actitudes reactivas y no reactivas sea asimilado al contraste entre relaciones interpersonales [actitud participante] y objetividad.” ⁵⁷³

En conclusión: el amor, los sentimientos heridos y cosas por el estilo forman parte de la actitud participante, sí, pero no son actitudes reactivas (a partir de ahora “emociones reactivas”). La actitud participante es aquella postura mental apta para la interacción social con personas normales y adultas.

⁵⁷⁰ De la que ya traté en *III.1.2.*

⁵⁷¹ Y cuando no lo hace así resulta forzado y extraño.

⁵⁷² Véase la discusión acerca del rol del terapeuta en las páginas 286, 310 y 317.

La actitud objetiva (que redefiní como actitud reflexiva en *II.4.1.*) es la postura mental apta para la investigación teleológica, que es una interacción excepcional con los demás, adoptada las más de las veces por ser aquellos con quienes interactuamos personas psicológicamente anormales o moralmente subdesarrolladas; pero no solamente, sino también cuando el contexto lo exige (en un juzgado, en un aula).

Resta ahora, tras estas puntualizaciones, tratar de averiguar de forma definitiva qué sentimientos son las emociones reactivas. Wallace suministra un buen criterio para distinguirlas. Pero poseer un buen criterio no implica que se tenga una buena taxonomía. Wallace cae en ese exceso de sobreintelectualización tan característico de la cuestión tratada. La borrosidad es sin duda un defecto de la teoría de Strawson, pero tal vez corregir en exceso este defecto es también acabar con una de sus virtudes. En el próximo apartado voy a intentar dar con un punto intermedio que ponga remedio al mismo tiempo a la excesiva borrosidad con que Strawson trata la cuestión y a la sobreintelectualización de los hechos en que incurre Wallace.

III.2.2. Hacia una taxonomía de las emociones reactivas

Según Wallace las emociones reactivas están siempre vinculadas a “la creencia de que cierta demanda ha sido violada”.⁵⁷³ Por ello en la teoría de Wallace desaparecen las emociones reactivas que podrían considerarse positivas en tanto que las sentimos, no cuando nuestra expectativa se ha visto frustrada, sino precisamente cuando se ha visto colmada. En todo momento Strawson habla de desaprobación pero también de aprobación moral; no así Wallace. En este sentido la lista de emociones reactivas de Wallace resulta algo

⁵⁷³ Wallace 1998: 29.

⁵⁷⁴ Ibidem: 24.

contraintuitiva, pues resentimiento, indignación y culpa tienen todos un carácter negativo, sin embargo parece que todos tenemos cierta idea de que puede atribuirse responsabilidad por acciones admirables, y no sólo por acciones reprobables. La postura cuasi evaluativa⁵⁷⁵ debería pues abarcar también la actitud proposicional en que se cree que una expectativa se ha visto cumplida, y no sólo cuando ha sido violada. ¿Es menos reactiva una emoción cuando es el correlato de la creencia de que cierta demanda ha sido cumplida? Mi idea es que Wallace no da ningún argumento satisfactorio para responder afirmativamente a esta pregunta. Así pues, querría proponer aquí una nueva taxonomía de las emociones reactivas, basada en la postura cuasi evaluativa pero teniendo en cuenta tanto las actitudes proposicionales en que se tiene la creencia de que cierta demanda ha sido violada, como aquellas en que se tiene la creencia de que ha sido cumplida. Se obtienen entonces algunas emociones reactivas positivas. Por otra parte, resulta útil a la hora de ordenar la clase de las emociones reactivas, además de fijarse en la postura cuasi evaluativa, aplicar un segundo criterio (olvidado por Wallace) que recojo directamente de Strawson: el de tener en cuenta hacia quién se vuelca la expectativa. No tendré en cuenta, en cambio, si el sujeto de las emociones reactivas participa de la acción o no. Strawson distingue entre las actitudes reactivas personales y las vicarias análogas basándose en ese criterio, pero sobre todo lo hace para poder distinguir entre las actitudes reactivas que él considera morales y las que no. Teniendo en cuenta que para esto maneja lo que, como se ha dejado patente, es un criterio equivocado (el de la vicariedad), no tendré en cuenta tal distinción.

⁵⁷⁵ Aquí tal vez el lector necesite refrescarse la memoria. Como expuse en las páginas 313 y 314, Wallace considera que estar en la postura cuasi-evaluativa es lo que Strawson llamó tener expectativas acerca de alguien, lo cual "(...) es esencialmente ser susceptible a cierto rango de emociones en el caso de que la expectativa no se vea colmada, o creer que la violación de la expectativa haría apropiado que uno fuese sujeto de esas emociones." Wallace 1998: 21.

Así, se obtienen dos grandes categorías: la de aquellas emociones en que considero que los demás han violado o cumplido mi expectativa; y la de aquellas en que considero que yo mismo he violado o cumplido mi expectativa. A las primeras, en tanto mi expectativa se vuelca hacia fuera, y por lo tanto mi emoción trata acerca de alguien diferente de mí, las llamaré *emociones reactivas intencionales*.⁵⁷⁶ A las segundas, en tanto mi expectativa se vuelve hacia dentro y por lo tanto mi emoción trata acerca de mí mismo, las llamaré *emociones reactivas autoreferenciales*. Ambas categorías se subdividen en razón de si las emociones reactivas aparecen por la creencia de que cierta expectativa ha sido violada, y en tal caso se hablará de *emociones reactivas negativas*; o si aparecen por la creencia de que cierta expectativa se ha visto cumplida, y en tal caso se hablará de *emociones reactivas positivas*. En el siguiente cuadro (véase la página siguiente) se distinguen mejor estas categorías, en él figuran además las emociones que se me aparecen más claramente como reactivas una vez analizadas a través de los criterios que acababo de señalar:

⁵⁷⁶ Se trata aquí de un uso metafórico del concepto de intencionalidad de Brentano. Considero que estas emociones son intencionales porque tratan sobre alguien diferente de uno mismo, y en ello son exactamente lo contrario de las autoreferenciales, que se vuelcan sobre el propio agente que las siente.

EMOCIONES REACTIVAS	Negativas	Positivas
Intencionales (“personales” y “análogas vicarias”)⁵⁷⁷	Resentimiento Indignación «Sentirse dañado u ofendido»	Gratitud Admiración Envidia
Autoreferenciales (“auto-reactivas”)	Culpa Vergüenza Remordimiento	Orgullo Autosatisfacción

No descarto que existan más, no es el objetivo de este trabajo construir una taxonomía completa de las emociones reactivas, pero considero que aquí figuran las más importantes. Naturalmente esta afirmación no puede hacerse a la ligera, por ello me voy a detener brevemente en analizar estas emociones reactivas a fin de justificar su presencia en el cuadro.

No voy a justificar incluir el *resentimiento*, la *indignación* y la *culpa*. Estas emociones reactivas han sido admitidas como tales por todos los autores que han sido estudiados hasta aquí en este trabajo y que se han ocupado de esta cuestión; no hago sino seguir la corriente dominante. Seguramente sí es necesario explicar el hecho de que resentimiento e indignación se encuentren en la misma categoría. Esta reunión de ambos conceptos se basa en el simple hecho de que Strawson se equivoca al pensar que el resentimiento sólo puede sentirse de forma personal y la indignación de forma vicaria: a menudo nos sentimos indignados por algo que se nos ha hecho a nosotros mismos y

también a menudo vemos que es posible sentir resentimiento por cosas que les han sido hechas a terceras personas. Sin ir más lejos sentimos resentimiento hacia los miembros del equipo de fútbol que elimina a nuestro equipo favorito en un mundial, o hacia alguien que le ha hecho un mal a un amigo. Y en cuanto a la indignación me sentiré, por ejemplo, más bien indignado que resentido si me hacen trampa en una oposición. Así pues, es un hecho que somos capaces de sentir resentimiento o indignación hacia los demás independientemente de nos hayan hecho un mal a nosotros mismos o a terceros. Por ello tanto el resentimiento como la indignación deben ser considerados igualmente emociones reactivas *intencionales* negativas.

En cuanto a la *gratitud* tampoco puede decirse que se trate de un caso especialmente ambiguo de emoción reactiva. Wallace es el único comentarista de Strawson que se niega a admitir la gratitud como emoción reactiva, pero si lo hace no es porque considere que la gratitud en particular no debería contar como tal, sino porque no está dispuesto a admitir ningún tipo de emoción reactiva de carácter positivo. Es así que he llegado ya a la conclusión de que al proceder de este modo Wallace sobreintelectualiza los hechos, y por lo tanto en lo que respecta a la gratitud sigo el ejemplo de Strawson y de la mayoría de sus comentaristas, y lo admito como una emoción reactiva intencional positiva.

Resulta extraño que ninguno de los filósofos que se han ocupado de esta cuestión haya tenido en cuenta esa emoción que no he conseguido nombrar más que con la expresión *sentirse dañado* (tal vez precisamente porque no hay, o al menos no se me ocurre, un nombre concreto para este sentimiento, nadie se ha ocupado de él). Precisaré que manejo aquí un concepto de daño que no incluye únicamente un daño físico, sino sobre todo un daño moral. Considero, por ejemplo, que una ofensa, como podría ser un

⁵⁷⁷ Entre paréntesis aparecen las categorías usadas por Strawson y encima de ellas mi categoría equivalente.

insulto, representa un mal que nos hace sentirnos dañados moralmente (sentirse ofendido es una variedad del sentirse dañado). Me parece que es un sentimiento que es difícil no admitir como un caso claro de emoción reactiva: precisamente consideramos una ofensa el que no se nos trate como se nos debe, esto es, el que se viole nuestra expectativa de ser tratados con “cierto grado de buena voluntad”, lo cual define, según Strawson, las emociones reactivas. El sentirse dañado u ofendido sobreviene cuando consideramos que ha sido violada una demanda de respeto. Me atrevería a decir que hay pocos casos tan claros como este de emoción reactiva. Es una lástima no contar con una denominación mejor para este sentimiento.

En lo que respecta a la *admiración*, hay que admitir que se trata de una emoción no exenta de ambigüedad, pues a veces aparece en casos en que no se puede decir que sea reactiva. Naturalmente el tipo de admiración al que me refiero aquí no es parecido a la idolatría enfermiza que pueda sentir un fan por una estrella del pop, sino más bien a la admiración que sentimos por la excelencia moral de una persona. A menudo sentimos admiración por una persona cuando vemos que, sin importarle las posibles consecuencias indeseables que pueda tener para él lo que hace, actúa de una cierta manera por el simple hecho de considerarlo su deber. En este sentido se puede decir que esa admiración se basa en el hecho de que creemos que alguien cumple cierta demanda que nosotros [*qua* sociedad] le hacemos. Se trata pues de la admiración que podemos sentir por Héctor cuando se despide de Andrómaca sobre las murallas de Troya: sabe que va a morir en la guerra pero considera que no puede renunciar a lo que es su deber, esto es, defender su ciudad contra el invasor. Dicho de otro modo: Héctor, a conciencia, decide que no puede violar la demanda de la ciudad que le pide protección, y por ello suscita nuestra admiración. En este sentido defiendo que la admiración es una emoción reactiva.

Sin duda choca ver de este lado, el de las emociones reactivas positivas, a la *envidia*, y más después de hablar de la admiración. Debo pues hacer aquí una observación: el que se juzgue que una emoción es positiva o negativa no tiene nada que ver con el hecho de que se trate de una emoción agradable o no, y sobre todo recomendable o no. Por ello no puedo dejar de considerar que la envidia, siempre que no se trate de un tipo enfermizo o compulsivo de envidia, como el que siente el Abel Sánchez de Unamuno, es una emoción de carácter positivo a pesar de que no parezca un sentimiento demasiado agradable. En efecto, salvo en aquellos casos en que sentimos envidia por objetos o cosas que nada tienen que ver con atribuciones de responsabilidad (como cuando envidiamos algo así como “la suerte” del otro), parece que la envidia puede llevar consigo una creencia de que cierta expectativa se ha visto cumplida. Por ejemplo, supongamos que alguien hace un bien a otra persona y que nosotros consideramos que ese bien debía ser hecho. Sin embargo, en vez de sentirnos alegres porque ese bien ha sido llevado a cabo, podemos sentir envidia porque nos vemos a nosotros mismos como capacitados para haber hecho ese mismo bien, que no pudimos hacer porque aquel a quien ahora envidiamos se nos adelantó. Estamos sintiendo envidia porque cierta demanda se ha visto cumplida. En cierto sentido *esta envidia es una admiración insana*. Así de extrañas son nuestras emociones: curiosamente hay un tipo de envidia que debe colocarse del lado de la admiración. E insisto en que nada tiene que ver el que una emoción se dé de forma positiva con el hecho de que sea considerada “reactiva positiva”: una demanda cumplida (que es lo que autoriza a considerar positiva a una emoción reactiva) puede provocar un sentimiento “nocivo”.

En lo que atañe al *orgullo*, éste debería entenderse también de forma similar a la admiración (y algo muy parecido puede decirse de la *autosatisfacción*, que tal vez no es sino una variedad del orgullo). Hay que verlo como esa sensación de deber cumplido, de trabajo bien hecho, como esa sensación placentera que nos sobreviene cuando hemos realizado algo meritorio, algún

logro del cual uno se siente plenamente responsable. No tiene nada que ver con el orgullo entendido peyorativamente que se usa irracionalmente en oraciones tan sinsentido como “me siento muy orgulloso de ser español”, como si pertenecer a cierta nación fuese algo de lo que pudiese uno sentirse gratamente responsable. Ahora, el que exista este orgullo irracional no obsta que haya una forma de entender el orgullo como una emoción reactiva autoreferencial positiva.⁵⁷⁸

La *vergüenza*⁵⁷⁹, al igual que otras emociones que ya he analizado, no está exenta de manifestaciones en que poco tiene que ver con atribuciones de responsabilidad: por ejemplo cuando sentimos vergüenza por cualquier rasgo físico que consideramos un defecto. Pero una vez más esto no significa que no exista un tipo de vergüenza que constituye una emoción reactiva autoreferencial similar al sentimiento de culpa, una vergüenza genuina que sentimos por haber hecho algo de lo cual nos sentimos responsables y que consideramos moralmente indecoroso, una violación de una demanda.

En cuanto al *remordimiento*⁵⁸⁰, desgraciadamente se trata de un concepto muy cargado de significado en nuestras sociedades cristianas y no siempre aquello a lo que se llama remordimiento puede considerarse una emoción reactiva: no cabe duda de que es posible tener remordimiento de forma plenamente irracional. Cuando hablo de remordimiento tengo en mente algo parecido a la indignación, sólo que volcado hacia uno mismo y con un deseo utópico de volver al pasado y corregir nuestra conducta. El hecho de que sea fácil confundir esta forma de darse el remordimiento, con otra que en nuestra

⁵⁷⁸ No puedo pasar por alto que a menudo sucede que alguien puede sentirse orgulloso de otra persona, así pues no se debería descartar la posibilidad de aceptar algo así como un “orgullo ajeno”, parecido a la admiración, como emoción reactiva intencional positiva.

⁵⁷⁹ Pasa con la vergüenza algo similar a lo que dice la nota anterior con respecto al orgullo. A veces podemos sentir algo a lo que le damos el nombre de “vergüenza ajena” que se podría considerar una emoción reactiva intencional negativa.

⁵⁸⁰ Todo lo que digo a continuación acerca del remordimiento considero que puede decirse igualmente acerca de la culpa (aunque la culpa parece tener una carga cognitiva mayor que el remordimiento).

sociedad es corriente sentir sin la capacidad de dar razones coherentes de por qué se siente, no elimina este otro remordimiento racional. Cuando hable aquí de remordimiento lo haré siempre como de un sentimiento de responsabilidad propia, no como el remordimiento insano del que se considera un pecador, un alma mancillada. El remordimiento cristiano no pinta nada en mi argumentación pues es, por definición, un remordimiento que nada tiene que ver con la responsabilidad en tanto es fruto de un pecado original, esto es, de una acción heredada de otros que sí fueron responsables de sus actos. Pero no todo remordimiento es el sentimiento que Nietzsche asimila a la culpa en el Tratado Segundo de su *Genealogía de la moral*.

Así pues, parece haber bastantes más emociones reactivas de lo que pretendía Wallace, pero seguramente no tantas como pensaban Strawson, Bennett y Wolf. Me inclino a pensar que las principales están en el cuadro que he elaborado, pero tal vez se les podrían sumar otras. Es cierto que algunas de las emociones que incluí son ambiguas, pero gracias a Wallace es suficiente con considerar en qué casos esas emociones constituyen una cuasi evaluación y en qué casos no, para saber cuando deben ser aceptadas como emociones reactivas. Es un grave prejuicio filosófico, de tintes dogmáticos, tender a aceptar dentro de una categoría sólo aquello que se puede inequívocamente admitir dentro de ella. Diagnosticar nuestros sentimientos nos resulta más fácil de lo que solemos pretender y basta dedicarle un mínimo de tiempo, como he hecho aquí, a hacer ciertas distinciones para lograr aclarar en qué consiste un determinado sentimiento, cuáles son sus diferentes manifestaciones y averiguar cuál de ellas seríamos capaces de admitir como emoción reactiva.

III.2.3. La indignación, la culpa y los límites de la moral

Más arriba se pudo ver que ni Bennett ni Wallace consiguen dar con un criterio plenamente satisfactorio de aquello que haría moral a una emoción reactiva. Con respecto a la solución de Bennett se vio que amplía demasiado los límites de la moral, de tal forma que prácticamente cualquier emoción reactiva podría ser considerada moral. En cuanto a Wallace se señaló que su solución al problema, si bien no plantea dificultades desde el punto de vista de su coherencia, no resulta satisfactoria porque simplemente desplaza el momento de explicar aquello que haría de algo que pueda ser llamado moral. Mi idea es que se puede combinar las posturas de estos dos autores para encontrar una solución satisfactoria a este problema.

Si, como acabo de recordar, el problema de Bennett es que toda emoción reactiva se convierte en sentimiento moral, el problema de Wallace parece ser exactamente el contrario. Si el que se pueda considerar que una emoción reactiva es moral depende del tipo de expectativa que la explique, entonces da la impresión de que no existirían emociones reactivas específicamente morales, sino más bien formas morales de darse las emociones, o justificaciones morales para las emociones, o situaciones morales en que se dan las emociones. La idea original de Strawson es que debe haber un tipo de emoción reactiva que sea específicamente moral, y esta idea no está exenta de importancia a la hora de justificar la existencia de la responsabilidad moral. En efecto, si solamente cabe distinguir emociones reactivas que señalan la existencia de responsabilidad, sin más, no es posible utilizar dichas emociones como argumento más que para justificar la presencia de cierto tipo de responsabilidad en nuestras vidas *que no es moral*.⁵⁸¹ Así, la perspectiva strawsoniana se convertiría simplemente en un arma útil para justificar una idea de la responsabilidad que realmente no tiene mucha importancia defender. Se debe, por lo tanto, si se aspira a presentar mediante las emociones

⁵⁸¹ Sobre la diferencia que existe entre una responsabilidad meramente causal y una responsabilidad moral trabajaré más detalladamente en *III.4.1*.

reactivas a la responsabilidad moral como un hecho, conjugar los sutiles análisis de Wallace con la intención de Strawson y Bennett de distinguir unas emociones reactivas que sean de carácter *esencialmente* moral.

Sin duda hay ciertas emociones reactivas que cabe considerar a veces son morales y a veces no, porque a veces se basan en expectativas que podemos considerar obligaciones y a veces no. Sin embargo, mi idea es que hay por lo menos dos emociones reactivas que siempre se basan en el hecho de que una obligación moral⁵⁸² ha sido violada: la indignación y la culpa. Desde mi punto de vista estas emociones cumplen siempre el requisito rawlsiano, en que se basa Wallace en su argumentación, para considerar que un sentimiento puede ser considerado moral. En efecto, según Rawls “es un aspecto necesario de los sentimientos morales, y parte de lo que los distingue de las actitudes naturales, que la explicación de su experiencia por parte de la persona invoque un concepto moral y sus principios asociados. La explicación de su sentimiento hace referencia a un bien o un mal reconocidos.”⁵⁸³ Así pues, lo que mantengo aquí es que la indignación y la culpa son unas emociones reactivas *morales* porque, siempre que se dan, son explicables por la creencia de que se ha violado cierta expectativa que no debía ser violada porque hacerlo constituye un mal desde el punto de vista, o bien de una teoría de la justicia, o bien de una teoría de la virtud, o bien de ambas. Así lo señala Rawls: “la explicación de algunos sentimientos morales se basa en principios del derecho que serían escogidos en la posición original, mientras que otros sentimientos morales están relacionados con el concepto de virtud”.⁵⁸⁴ Voy a considerar que indignación y culpa son emociones reactivas morales porque, siempre que se dan, lo hacen porque ciertas expectativas que se justifican mediante principios han sido violadas. No comparto la idea de que deba existir

⁵⁸² Utilizo aquí “moral” en sentido amplio (puede abarcar la política).

⁵⁸³ Rawls 1971 [1999]: 421.

⁵⁸⁴ Ibidem: 422.

necesariamente una dicotomía entre la teoría que nos permite distinguir entre lo bueno y lo malo, y la que nos permite hablar de lo justo y lo injusto.⁵⁸⁵ Aquí voy a hablar de justicia y de virtud como dos caras de una misma moneda: de un sistema de valores.⁵⁸⁶ No me interesan realmente los pormenores de los conceptos de justicia y de virtud, tan sólo la forma en que funcionan como estructura y sustento de la moral, y por lo tanto como elementos que dotan de carácter moral a las emociones reactivas que he señalado como específicamente morales. Cuando sentimos indignación, *siempre* es porque consideramos que ha sucedido algo que es malo o injusto, y lo mismo puede decirse de la culpa.

En cuanto a la indignación no parece que pueda haber muchos problemas para admitir esto. Esa relación de la indignación con una teoría de la virtud (que pretende sustentarse en principios objetivos sobre lo bueno y lo malo) la explica concienzudamente Antonio Valdecantos: “Quien se indigna tiende a hacer propaganda de su indignación, y el mejor medio de propaganda es aportar las razones de la emoción. Está persuadido de que sus razones son buenas y esto significa que son *suyas* de manera puramente accidental; igualmente podrían ser de cualquier otro (...).”⁵⁸⁷ En efecto, la indignación es una emoción que consideramos debería ser sentida por cualquiera que conociese el caso que nos empuja a nosotros a sentirla. Si hay razones para ella entonces no puede no ser sentida: “(...) cuando estoy indignado por algo y expongo a alguien las razones de mi indignación, no sólo espero que las comparta, sino también que de ellas se siga en mi interlocutor un estado de

⁵⁸⁵ No puedo sino reconocer que aquí estoy expresando una idea que no soy capaz de justificar adecuadamente, sobre un tema que además no es baladí. Probablemente necesitaría otra tesis para justificar esta intuición, para quien esté interesado en el debate que conlleva me remito a una nota futura, la 611.

⁵⁸⁶ En *III.4.1.* y *III.4.3.* me ocuparé de explicar en qué consiste un sistema de valores y cómo se adquiere.

⁵⁸⁷ Valdecantos 2001: 80.

ánimo semejante al mío.”⁵⁸⁸ Esto no pasa con otras emociones reactivas: si estoy resentido y explico a alguien las razones de mi resentimiento no le exigiré que él a su vez sienta resentimiento (aunque tal vez me agradaría que así fuese), sino tan sólo que comprenda mis razones y tal vez me dé su aprobación o me deje patente que cree que mi resentimiento está justificado. Por el contrario el que yo tenga razones para sentirme indignado obliga a cualquiera a quien le exponga dichas razones a sentirse indignado a su vez. En este sentido la indignación está directamente ligada a la justicia: espero de alguien que me diga “esto es indignante” en el mismo sentido en que podría decirme “esto es una injusticia”. Volviendo al ejemplo del fútbol: sentiré resentimiento hacia el delantero de mi equipo que nos hizo perder por fallar un penalti, pero sentiré indignación hacia el árbitro por habernos hecho perder anulándonos un gol legal. En el primer caso siento resentimiento porque la violación de la expectativa que yo tenía no es una violación de la justicia, en el segundo caso siento indignación porque sí lo es. Así pues hay que admitir a la indignación como una emoción reactiva moral *per se*.

En lo que respecta a la culpa es cierto que es una emoción que resulta más ambigua, no obstante se puede afirmar que en la culpa hay algo esencialmente moral y por ello nos consuelan expresiones como “tú no hiciste nada malo” o “no fue tu intención” o “hiciste lo que pudiste”. Si atropello a alguien porque me desmayo porque sufro una embolia mientras conduzco, seguramente me sentiré mal por la muerte de esa persona y probablemente también me sentiré responsable. Y lo seré, pero sólo físicamente, causalmente, y no moralmente, luego no me sentiré culpable: el que me sienta responsable no implica que me sienta culpable, hay un juicio moral implícito en el sentimiento de culpabilidad, y si no lo hay entonces no es una culpabilidad genuina. Así lo demuestra Rawls con el siguiente ejemplo: “pueden haber

⁵⁸⁸ Ibidem.

enseñado a una persona educada en una estricta secta religiosa que ir al teatro es malo. Cuando ya no cree esto, nos dice que todavía se siente culpable cuando acude al teatro. Pero estos no son sentimientos de culpabilidad genuinos, ya que no tiene intención de disculparse ante nadie, o de decidir no ver ninguna otra obra, o algo así. En vez de ello dirá más bien que tiene ciertas sensaciones y sentimientos de inquietud, o algo así, que se parecen a aquellos que uno tiene cuando se siente culpable.”⁵⁸⁹

¿No hay sentimientos positivos específicamente morales? Estoy prácticamente convencido de que sí, pero no me atrevo a asegurarlo tan nítidamente como considero puedo hacerlo en el caso de la indignación y la culpa. El caso es que nuestra sociedad pone el peso del juicio moral en la condena más que en el premio, con lo que los conceptos que suelen llevar asociada una importante carga moral suelen ser de carácter negativo y no positivo. Se tiende a considerar que aquello que sería susceptible de provocar emociones reactivas positivas caería del lado de lo supererogatorio, y como nos hemos hecho todos muy kantianos tendemos a considerar que es moral lo que se hace por deber. Sospecho que la gratitud podría ser considerada una emoción reactiva específicamente moral, pero no poseo argumentos para demostrarlo más allá de la intuición. También parece que la admiración podría serlo, pero no cualquier tipo de admiración, sino una clase “trágica” o “aristotélica”, no muy corriente hoy en día.⁵⁹⁰ En Grecia no me cabría la menor duda a la hora de asignarle un valor moral a la gratitud y a la admiración: recuérdese que el bien para los griegos es hacer el bien a los amigos y el mal a los enemigos, luego sentir gratitud por un bien que se nos hace es algo moralmente obligatorio; por otra parte la admiración típicamente griega es aquella en que me centré cuando traté específicamente de este sentimiento, la que uno ha de sentir por Héctor defendiendo Troya, y que se

⁵⁸⁹ Rawls 1971 [1999]: 422.

⁵⁹⁰ Véase a este respecto Arteta 2002.

distingue de formas espúreas de admiración por su carácter más moral, más bien basado en el valor y la bondad de la acción que en su espectacularidad (la teoría aristotélica de la tragedia es el mejor instrumento que hay para comprender este punto). Así, creo que el cuadro de las emociones reactivas morales debería incluir a la indignación y la culpa del lado negativo, y a la gratitud y la admiración del lado positivo.

¿Por qué no incluir en esa lista el resentimiento, al que Wallace considera moral? Porque no parece estar siempre vinculado con una teoría de la justicia o de la virtud. He aquí un ejemplo: estoy jugando al fútbol (una vez más) con mis amigos, mi equipo y el contrario están empatando y en el minuto noventa se pita un penalti a nuestro favor, lo tira nuestro delantero y lo falla, finalmente acabamos perdiendo el partido en la prórroga. En este caso puedo sentir resentimiento hacia mi delantero sin ninguna conexión con la justicia, sino sencillamente porque se ha fallado en un juego. El resentimiento no parece ser específicamente moral.

En fin, no puedo dejar este apartado sin señalar que, como bien dice Wallace, seguramente todas las emociones reactivas pueden ser morales en algún momento porque hacen referencia a obligaciones morales, pero eso no las convierte en emociones morales en el sentido en que las he tratado aquí, esto es, como emociones *esencialmente* morales. No obstante es importante no cerrar esa puerta. Existen dos emociones reactivas específicamente morales, la indignación y la culpa, y otras dos más susceptibles de serlo, la gratitud y la admiración. Todas las demás pueden ser morales cuando se apoyan en expectativas que convendríamos en considerar obligaciones. En potencia toda emoción reactiva es moral, necesariamente lo son la indignación y la culpa, y (dicho con la boca pequeña) la gratitud y la admiración.

III-3. El arte de excusar: ¿Qué vale y qué no vale?

“Como ‘real’, ‘libre’ sólo se usa para excluir la sugerencia de algunas o todas sus antítesis reconocidas. Así como ‘verdad’ no es el nombre de una característica de las aserciones, sino el nombre de una dimensión en la que las acciones son juzgadas. Examinando todas las formas en que una acción pueda no ser ‘libre’, i.e. los casos en que no sirve decir simplemente ‘X hizo A’, podemos esperar deshacernos del problema de la Libertad.”⁵⁹¹

En la parte II del trabajo traté de mostrar que Strawson comete un grave error al presuponer que las emociones reactivas, por pertenecer al terreno de los sentimientos, están fuera de todo fundamento racional. Ese error le conduce a la conclusión de que nunca nada podría hacerlas desaparecer, lo cual, como se vio entonces imaginando un mundo determinista de “leyes manejables”, es claramente falso. Tenemos la posibilidad de modificar nuestras emociones reactivas en razón de si las consideramos o no justificadas racionalmente. De hecho esa posibilidad es la que hace que seamos capaces de oscilar entre la actitud que terminé llamando reactiva y aquella que distinguí como reflexiva. Dicho de otro modo: somos capaces de eximir o excusar de responsabilidad dentro de nuestra vida participante en razón de que únicamente sentimos las diferentes emociones reactivas implicadas en nuestro juego de atribución de responsabilidad cuando consideramos que están justificadas racionalmente. Erróneamente se tiende a pensar únicamente en la naturaleza cuando se habla de sentimientos. Se asocian los sentimientos al cuerpo y se tiende a considerar que éste es algo plenamente natural, sin embargo, si se tienen en cuenta las enseñanzas de algunas feministas: no sólo el género es un constructo cultural, también el sexo, en parte, lo es. Así, habría que admitir que también nuestra forma de ver nuestro cuerpo depende en gran medida de nuestra cultura, y más aún los sentimientos que le dan vida. Por ello voy a estudiar a continuación los anclajes culturales de nuestras emociones

reactivas y las estructuras lingüísticas que dotan de sentido a nuestro juego de excusas y exenciones.

III.3.1. Emociones culturales y actitud reflexiva

Voy a tratar de corregir aquí un peligroso error filosófico al que empuja cierto tipo de naturalismo, que acaba siendo una forma de esencialismo. Se trata de una interpretación internista del hecho de que nos consideramos seres responsables. Así, Thomas Nagel en su artículo “Moral Luck” dice lo siguiente: “Hay una conexión cercana entre nuestros sentimientos acerca de nosotros mismos y nuestros sentimientos acerca de otros. Culpa e indignación, vergüenza y desdén, orgullo y admiración son caras internas y externas de las mismas actitudes morales. [...] Aplicamos la misma concepción esencialmente interna del yo a los demás.”⁵⁹² Y en un tono similar Galen Strawson dice en su *Freedom and Belief*: “(...) parece posible que un ser pueda desarrollar un fuerte sentido de la libertad (un sentido de la libertad que sea en todo lo esencial el mismo que el nuestro) en un mundo en el que no hubiera otras criaturas como él y criaturas que fueran tales que le llevaran a suponer que eran libres (...).”⁵⁹³ O incluso: “Ninguno de esos sentidos de la libertad depende esencialmente de la interacción con otra gente, o de la preparación para atribuir libertad a otros, sino solamente en la experiencia privada de la deliberación y la acción.”⁵⁹⁴

Sería un error negar la evidencia de que tenemos conciencia de nuestras acciones y nuestras deliberaciones; no es ahí, en el hecho mismo de nuestra autopercepción como seres libres, donde está el debate. Las dificultades vienen a la hora de señalar el origen de esa conciencia de la libertad. Los testimonios

⁵⁹¹ Austin 1956-7, en Urmson & Warnock 1961: 180.

⁵⁹² Nagel 1979, en Watson 1982: 184-185.

⁵⁹³ Strawson, Galen, 1986: 114.

⁵⁹⁴ *Ibidem*: 115-116.

que acabo de citar se inclinan por la idea de que se trata de una experiencia privada que no necesita de la presencia de otros seres para poder ser desarrollada. Desde mi punto de vista para sostener esta hipótesis internista es necesario adoptar una postura esencialista. En efecto, sólo si se postula un tipo de libertad esencialmente ligado a la naturaleza humana e independiente de cualquier tipo de factor social se puede creer que puede sobrevenir la experiencia de la responsabilidad fuera de la sociedad. A este punto de vista se le puede hacer las objeciones que he venido haciendo ya a la estrategia naturalista “dura” a lo largo de este trabajo. Parece difícil de creer que un ser humano criado en estado salvaje y aislado socialmente pueda llegar a una experiencia efectiva de la responsabilidad más allá de la misma experiencia que puedan llegar a tener de ella un perro o un gato, a saber: la de que sus miembros obedecen a sus deseos. Por el contrario todo empuja a sostener que nuestra percepción de la responsabilidad es adquirida y por lo tanto cultural. Parece necesaria cierta configuración social que dote de sentido a la responsabilidad,⁵⁹⁵ y no se puede descartar la idea de que haya culturas que funcionen sin concepto de responsabilidad o con conceptos afines pero no comparables (porque esto es posible funciona el argumento de Ayer acerca de un mundo determinista de “leyes manejables”).⁵⁹⁶ ¿Cuáles son las consecuencias de adoptar este punto de vista?

La primera de estas consecuencias es que nuestras emociones reactivas están, a pesar de ser emociones, muy lejos de ser desnudas expresiones de nuestra naturaleza. Nuestras emociones reactivas tienen una gran parte de carga cultural. Sin duda el componente emotivo se deriva de nuestros sentimientos más básicos (alegría, tristeza, ira, miedo, sorpresa y asco), no

⁵⁹⁵ El apartado III.4. trata esencialmente de esa configuración social.

⁵⁹⁶ Sobre culturas que basan su comunidad moral en conceptos diferentes del de responsabilidad (en concreto en el de vergüenza) consúltense Wallace 1996: 31-32, 38-40, 43-44 y 65-66; y Gibbard 1990.

obstante se trata de emociones fruto de una mediación cultural muy grande y de un complejo juego de significados que ordenan nuestra comunidad moral.

La segunda consecuencia de una visión pragmatista, fuertemente normativista, de la responsabilidad moral nos convierte a todos en sujetos de una comunidad moral regida por ciertos principios que ordenan nuestro juego de expectativas en que se basan nuestras emociones reactivas. Así se puede decir, en terminología Wittgensteiniana, que la responsabilidad moral es algo así como el “trasfondo” de nuestras sociedades, un trasfondo que nos permite entendernos en nuestra compleja red de relaciones participantes y que nos convierte en agentes que dentro de esa red tienen capacidad para formular expectativas y para ser requeridos para dar cuenta de la expectativas de los demás.

En el apartado anterior, en que me he ocupado de las emociones reactivas, he llegado a la conclusión de que cada una de ellas constituye en sí misma un juicio sobre expectativas cumplidas o incumplidas. He llegado también a la conclusión de que algunas de estas emociones reactivas (la indignación y la culpa, y tal vez la gratitud y la admiración) son específicamente morales, si bien todas ellas pueden serlo en razón del tipo de expectativa que incluyan. Pues bien, todo este entramado que he decidido llamar el mundo participante, y que no es sino una sofisticada estructura cultural de fondo, constituye lo que de forma harto imprecisa llama Strawson en su artículo una “comunidad moral” y cuyo significado intentaré aclarar más adelante. Ahora voy a ocuparme del problema siguiente: si el compromiso fundamental con nuestro mundo participante no es natural, sino que se trata de un compromiso social básico, ¿cómo se debe entender lo que he llamado “la actitud reflexiva”? En efecto, Strawson distingue entre la actitud participante (ahora actitud reactiva) y la actitud objetiva (ahora actitud reflexiva), porque considera que la primera se rige por reacciones naturales que conforman nuestras relaciones interpersonales, mientras que la segunda

consiste en el abandono de esta forma natural de relación y su sustitución por una, en palabras de Jonathan Bennett, “investigación teleológica”. Ahora nos encontramos con que ambas actitudes son fruto de un compromiso sociolingüístico y por lo tanto no se las puede distinguir por su grado de “naturalidad”. Además, en III-1. se convino en que las emociones reactivas que constituyen la actitud reactiva tienen un claro carácter cognitivo, ni mayor ni menor que el que pueda tener la actitud reflexiva: ambas actitudes forman parte del mundo participante. Así pues, si no se puede señalar una diferencia esencial entre ambas actitudes, no se puede por menos que pensar que lo que las separa es la *forma* en que se dan. La actitud reactiva y la actitud reflexiva no son sino diferentes formas de darse la vida participante. Dependiendo del sujeto que tomemos en cuenta, y del contexto que nos rodee, pensaremos que lo apropiado es comportarnos reactivamente o por el contrario comportarnos reflexivamente. Así, es apropiado, por ejemplo, comportarse reflexivamente en un juzgado (si se es miembro del jurado) o en un aula (si se es profesor). Hay contextos que hacen obligatoria la actitud reflexiva. Entonces es cuando se dan lo que llamé “veredictos” (atribuciones reflexivas de responsabilidad).⁵⁹⁷

Así, si se acepta que las emociones reactivas son un producto cultural, se puede resolver lo que más arriba consideré era una contradicción en la teoría de Strawson: la actitud reflexiva, adecuada para el niño o el enfermo, es a menudo adoptada para con la gente normal. Esto es posible porque, una vez se dota a las emociones reactivas de un componente cognitivo cuyo fundamento es social, éstas ya no aparecen como reacciones naturales inevitables, sino como reacciones de las cuales somos dueños en gran medida gracias a la razón. Por ello podemos ser capaces de abandonarlas si eso es lo que nos interesa en un momento dado, independientemente de que nos encontremos o no ante alguien “incapacitado para las relaciones

⁵⁹⁷ Véase la página 288.

interpersonales ordinarias”. Racionalmente podemos considerar que, de forma excepcional, es aconsejable que adoptemos la actitud reflexiva en una ocasión en la que tenderíamos a pensar que lo apropiado es adoptar la actitud reactiva.

En cualquier caso, desde el momento en que se decide vincular las emociones reactivas a ciertas actitudes proposicionales con forma de expectativa, propias de la postura cuasi evaluativa, hay “que suponer que [las emociones reactivas] tienen que tener una dimensión cognitiva: en particular, parece que una persona objeto de una actitud reactiva debe tener cierto tipo de creencia evaluativa, una que figure en la explicación de su estado emocional”.⁵⁹⁸ No puedo pues por menos que concluir que nuestras emociones reactivas están basadas en un complejo entramado cultural que compone nuestro mundo participante, y que consta de ciertas normas que nos indican cuando es apropiada la actitud reactiva y cuando, por el contrario, ésta debe ceder su sitio a la actitud reflexiva. Ahora me toca ocuparme precisamente de las normas en que se asienta nuestro mundo participante, y más concretamente de construir una taxonomía de aquello que convencionalmente estamos dispuestos a admitir como excusas o exenciones.

III.3.2. Excusas y exenciones

La existencia de un sistema sociolingüístico de excusas y exenciones es uno de los mejores argumentos que cabe dar en contra del incompatibilismo (una vez se ha mostrado que el incompatibilista está fuera del mundo participante). Siguiendo la perspectiva wittgensteiniana adoptada en II.4., se puede decir que la creencia en el determinismo y la convicción incompatibilista, una vez formuladas, no son sino excusas sin sentido, constituyen un mal uso del lenguaje o un síntoma de que no se entiende bien

el mismo: “Sólo el contexto habitual hace posible que se muestre con claridad lo que se quiere decir.”⁵⁹⁹

En efecto, imagénese una conversación en que un incompatibilista trata de convencer a sus vecinos de que no es responsable de haber atropellado a su perro porque sencillamente la libertad no existe y por lo tanto él no es responsable de nada. Seguramente lo primero que pensarán los vecinos del canicida incompatibilista al decirles éste “No soy responsable de haber atropellado a su perro” será que no fue realmente él quien lo atropelló, sino otra persona. Pero entonces nuestro incompatibilista dirá: “No hombre. Por supuesto que lo atropellé yo. Pero no soy responsable de ello.” Con un ligero estupor los dueños del malogrado animal dirán entonces: “¿Se cruzó en su camino sin que usted pudiese esquivarlo?”, o bien “¿No tuvo más remedio que atropellarlo para esquivar a un niño imprudente que se metió en la carretera?”, o tal vez “¿Tuvo un ligero desmayo, fue incapaz de verlo?”. En fin, los vecinos buscarán cualquier tipo de excusa realmente válida, esto es, con sentido en su lenguaje, en su mundo (que es el nuestro). Pero el incompatibilista se empeñará en dar la única excusa que no tiene ningún sentido en ese (este) lenguaje: “No. No soy responsable del atropello porque no pude obrar de otro modo, y esto porque nunca puedo obrar de otro modo que como obro efectivamente, porque nunca soy responsable de mis actos.” Los vecinos entonces dudarán entre llamar directamente a un manicomio o seguir con esa estúpida conversación, pero como hasta ese momento su vecino, dejando al margen el hecho de que era filósofo, les había parecido alguien muy normal, con envidiable buena fe seguirán preguntando: “¿Toma usted medicación? ¿Está usted enfermo y no sabe lo que hace?” Llegado ese punto sin duda nuestro incompatibilista se indignará y dirá vehementemente: “¡No entienden ustedes nada! Nadie es responsable de nada, no existe la libertad.” A partir de

⁵⁹⁸ Wallace 1998: 33.

⁵⁹⁹ Wittgenstein 1988, § 237: 31c.

aquí podrían suceder dos cosas: o bien los dueños del difunto perro, que leen la *Ética* de Spinoza en sus ratos libres, recordarán que todos estamos determinados y no somos metafísicamente libres; o bien llamarán a la policía. El resultado, antes o después, de cualquiera de estos dos hechos acabará siendo el mismo: nunca el incompatibilista llegará a ser capaz de usar el incompatibilismo como excusa que le permita eludir su responsabilidad. El intentar hacerlo es jugar mal el juego en el que todos estamos inmersos, es no participar del mismo mundo. Lo lógico sería pensar que el incompatibilista no está en sus cabales o que no comprende bien el idioma. Tal vez los vecinos, grandes lectores de Spinoza como ya sabemos, disfrutarían un momento discutiendo sobre metafísica con nuestro incombustible incompatibilista, pero a la larga no dejarían de juzgarle responsable de la muerte de su perro, aunque es muy posible que al final le dejaran ir en paz por considerarle un lunático. Lo más probable es que acabasen rayándole el coche y mudándose.

Esta es una historia divertida (si se obvia la muerte del pobre perro), pero el problema que plantea es un problema serio. Nuestro sistema de excusas y exenciones requiere una competencia interpretativa de carácter lingüístico y “ese proceso interpretativo asume que las acciones de un agente tienen lugar dentro de un medio social. Un agente debe comprender que dentro de ese medio social, a una acción-tipo se le asigna convencionalmente un significado *normalmente* indicativo de cierto tipo de motivo. Esto no es sugerir que cada uno deba conformar sus acciones particulares a una interpretación asignada a alguna acción-tipo. Es sólo sugerir que si nuestras acciones han de divergir de esas interpretaciones, sus variaciones deben de alguna manera ser una función del entramado interpretativo en que operemos.”⁶⁰⁰ El determinismo metafísico no es una función dentro de ese entramado y por ello el incompatibilismo no funciona como una excusa. Está

⁶⁰⁰ McKenna 1998b: 4.

directamente fuera del mundo participante, para cobrar sentido necesita de un mundo diferente, de otro entramado interpretativo. “En el contexto del día a día, cuando legisladores, jueces, padres y demás gente se ven concernidos por resoluciones de responsabilidad, saben, o creen que saben, cuáles son en general las condiciones de la responsabilidad. Sus preguntas son cuestiones de aplicación (...). Se asume, en esos contextos, que los humanos adultos totalmente desarrollados y normales son seres responsables. Sus preguntas tienen que ver con si un individuo dado cae dentro del rango de lo normal.”⁶⁰¹

Es indispensable, pues, un catálogo de exenciones y excusas que convencionalmente consideremos válidas para desarrollar nuestra vida participante. Por supuesto no se trata de un catálogo explícito, de hecho se trata de un catálogo tan impreciso que en parte debido a esa vaguedad existe el sistema judicial: para permitirnos decidir en cada caso si hay o no algo que excusa o exime de responsabilidad al presunto culpable de turno. No obstante hay ciertos rasgos generales de las excusas y exenciones que se pueden llegar a identificar. Sin ir más lejos con gran maestría Strawson logró una taxonomía más que correcta, distinguiendo tipos de expresiones que solemos usar para eludir nuestra responsabilidad. En sus categorías me voy a basar para tratar de desarrollar una taxonomía útil que más adelante me permitirá distinguir entre dos caras de la responsabilidad que dan cuenta de la complejidad de nuestra vida participante y del fenómeno de la atribución de responsabilidad en general.

En *II.1.1.*, se vio que hay dos grandes grupos de circunstancias que nos conducen a adoptar una actitud reflexiva (en Strawson “objetiva”): las circunstancias que excusan y las circunstancias que eximen. Las excusas no nos dicen nada sobre un agente, sino sobre una acción. Nos dicen que por *esa* acción, aunque tal vez fuera llevada a cabo voluntariamente, no podemos

⁶⁰¹ Wolf 1988, en Kane 2002: 146.

responsabilizar al agente. Como ya se dijo, las razones que se le ocurren a Strawson para que esto suceda no son diferentes de las que presenta Aristóteles en su *Ética nicomáquea*, a saber: la ignorancia y la fuerza mayor. Excusamos a un agente por una determinada acción cuando consideramos que no tuvo otra opción porque alguien o algo le obligó a actuar del modo en que lo hizo. Ahora, lo que es propio del agente para el que las excusas son adecuadas es que es dueño de su voluntad: excusamos de responsabilidad a un agente cuando actúa voluntariamente, pero en aquellas ocasiones en que su voluntad fue forzada por algo ajeno a él. Por ejemplo: un demente que odia las casas pintadas de verde me obliga, a punta de pistola, a incendiar el precioso chalé verde de mi vecina, que por otra parte disfruta plácidamente de sus vacaciones con su familia fuera del país. Naturalmente yo incendio el chalé de mi vecina. Pues bien, es obvio que actuó voluntariamente, podría elegir no quemar el chalé y que el demente me volase la tapa de los sesos; pero no se trata de una posibilidad real de elección, pues una posibilidad es claramente menos halagüeña que la otra y yo no tengo ningún medio a mi alcance para cambiar esas posibilidades. Así, la pistola en mi sien es una buena excusa que me convierte en un objeto inadecuado para el resentimiento de mi vecina. En cualquier caso, lo que es propio de las excusas es que sólo pueden ser aplicadas a alguien que normalmente es dueño de sí mismo, un sujeto normalmente libre y responsable que se ve forzado a actuar de una cierta manera sin que por ello pueda decirse que al hacerlo no actúe según su voluntad, que haya perdido el control.

En lo que respecta a las exenciones sucede todo lo contrario, para poder ser eximido de responsabilidad es indispensable actuar no siendo dueño de uno mismo: sin voluntad o con una voluntad compulsiva. Como se vio en II.1. esto puede suceder de dos maneras: o bien porque uno pierde transitoriamente su dominio sobre sí mismo, o bien porque no se tiene nunca ese dominio, porque sencillamente se es así.

Para resumir más claramente las caracterizaciones que acabo de hacer, voy a considerar que aquello que he llamado “ser dueño de uno mismo” o “ser capaz de actuar voluntariamente” es lo mismo que “pertenecer a la comunidad moral” y, al contrario, “no ser dueño de uno mismo” o “ser incapaz de actuar voluntariamente” es lo mismo que “no pertenecer a la comunidad moral”. Más adelante investigaré con más calma en qué consiste pertenecer o no a la comunidad moral, pero por manejar unos términos que ya son familiares se puede decir que es aquello a lo que Strawson llama “estar capacitado para las relaciones interpersonales normales”. En cualquier caso se obtiene que las excusas se aplican a los sujetos que pertenecen a la comunidad moral mientras que las exenciones se aplican a aquellos que, al menos transitoriamente, no pertenecen a la comunidad moral.

A los *miembros de la comunidad moral* les aplicamos las expresiones que señala Strawson como ejemplos típicos de excusa: “expresiones tales como ‘No pretendió’, ‘No se dio cuenta’, ‘No sabía’; o también todas aquellas que podrían darnos ocasión de usar la frase ‘No pudo evitarlo’, cuando se apoya en frases tales como ‘Le obligaron’, ‘Fue la única salida’, ‘No le dejaron otra alternativa’, etc.”⁶⁰²

A los que están *fuera de nuestra comunidad moral* les aplicamos expresiones como “‘No es más que un niño’, ‘Es un esquizofrénico incurable’, ‘Su mente ha sido metódicamente corrompida’, ‘Eso es un comportamiento puramente compulsivo por su parte’.”⁶⁰³ Y a aquellos que son juzgados por algo que hicieron estando transitoriamente fuera de nuestra comunidad moral les aplicamos expresiones “tales como ‘No era él mismo’, ‘Ha estado soportando una gran tensión recientemente’, ‘Estaba actuando bajo sugestión

⁶⁰² Strawson, P. F. 1962, en Fischer & Ravizza 1993: 50.

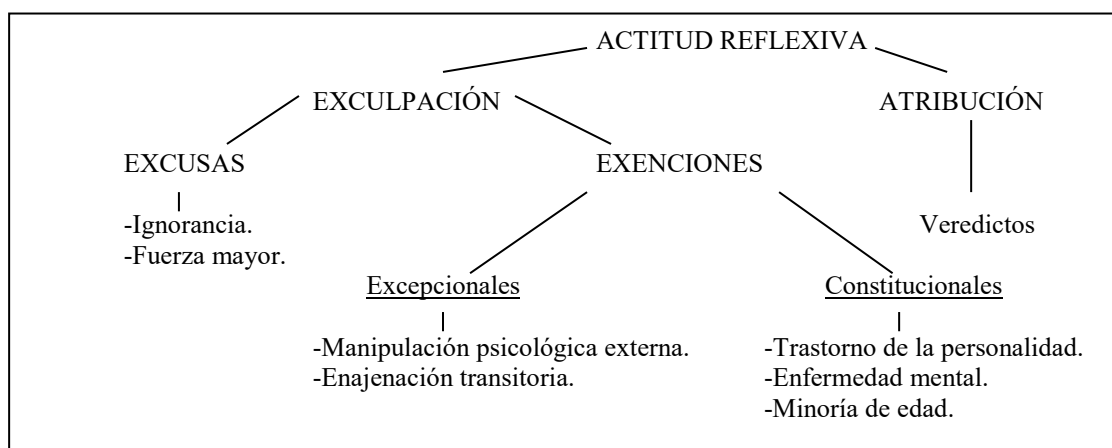
⁶⁰³ Ibidem.

posthipnótica’.”⁶⁰⁴ Estos dos últimos grupos de expresiones, también señalados por Strawson, son buenos ejemplos de *exenciones*.

Por el momento no puedo definir más exactamente la pertenencia a la comunidad moral, pero entendiéndola intuitivamente, como vengo haciéndolo en estas últimas páginas, y gracias a los ejemplos que suministra Strawson, es posible señalar:

- cuáles son las principales razones (excusas) para obviar la responsabilidad moral de alguien que es miembro de nuestra comunidad moral.
- cuáles son las principales formas de dejar de pertenecer a la comunidad moral (exenciones), tanto transitoriamente (*exenciones excepcionales*) como permanentemente (*exenciones constitucionales*), y por lo tanto de no ser tenido por responsable moralmente.

Pues bien, con esto ya se puede elaborar una taxonomía de las excusas y las exenciones que conforman parte de la actitud reflexiva:



En cuanto a la *excusa de la fuerza mayor* me remito directamente a Aristóteles:

⁶⁰⁴ Ibidem: 51.

[E]s forzoso aquello cuyo principio es externo y de tal clase de que en él no participa ni el agente ni el paciente; por ejemplo, si uno es llevado por el viento o por hombres que nos tienen en su poder. En cuanto a lo que se hace por temor a males mayores o por alguna causa noble (por ejemplo, si un tirano que es dueño de los padres e hijos del agluen mandara a éste hacer algo vergonzoso, amenazándole con matarlos si no lo hacía, pero salvarlos si lo hacía), es dudoso si este acto es voluntario o involuntario. Algo semejante ocurre cuando se arroja el cargamento al mar en las tempestades, nadie sin más lo hace con agrado, sino que por su propia salvación y la de los demás lo hacen todos los sensatos.⁶⁰⁵

También Aristóteles explica suficientemente bien en qué consiste la *excusa de la ignorancia*:

Todo lo que se hace por ignorancia es no voluntario, pero, si causa dolor y pesar, es involuntario. [...] [P]uede uno ignorar lo que hace, por ejemplo, cuando alguien dice que se le escapó una palabra o que no sabía que era un secreto, como Esquilo con los misterios, o que, queriendo mostrar sólo su funcionamiento, se le disparó como el de la catapulta. También podría uno creer que su propio hijo es un enemigo, como Mérope; o que la punta de hierro de la lanza tenía un botón; o que una piedra cualquiera era piedra pómez; o dando una bebida a alguien para salvarlo, matarlo por el contrario; o queriendo a uno darle una palmadita, noquearlo como en el pugilato.⁶⁰⁶

En lo que respecta a las *exenciones*, las *excepcionales* se caracterizan porque uno deja de ser sí mismo transitoriamente y, como decía Strawson “no deberíamos sentir resentimiento hacia una persona por una acción llevada cabo por la persona que no es.”⁶⁰⁷ Bajo la categoría de *manipulación psicológica externa* caen todas aquellas ocasiones en que la enajenación transitoria se debe a la acción manipuladora de un agente externo como cuando se actúa bajo sugestión posthipnótica, o como efecto de un lavado de cerebro... Bajo la categoría de *enajenación transitoria* caen aquellas ocasiones en que se actúa sin ser uno mismo por el efecto de algún tipo de droga que uno mismo se ha administrado, o por estar soportando situaciones de grave tensión... En fin, en situaciones en que un factor externo que no es una persona que desea utilizarnos no nos permite actuar de acuerdo con nuestra voluntad.

Por último las categorías que componen las *exenciones constitucionales* son clásicas. Considero *enfermedad mental* a trastornos psicológicos graves que nos

⁶⁰⁵ Aristóteles, *Op. Cit.*, Libro III (1110a).

⁶⁰⁶ Ibidem, (1110b 19) y (1111a 7).

⁶⁰⁷ Strawson, P. F. 1962, en Fischer & Ravizza 1993: 51.

incapacitan para conocer el alcance real de nuestras acciones, que distorsionan nuestra percepción y que hacen imposible que tengamos un funcionamiento cognitivo normal (por ejemplo una esquizofrenia). Dichas dolencias son inevitables, fruto de una predisposición, si no predeterminación, genética altamente poderosa. Considero *trastorno de la personalidad* a un trastorno psicológico que no supone ninguna mengua de nuestro funcionamiento mental, sino de nuestras capacidades sociales (por ejemplo una psicopatía o un trastorno de la personalidad antisocial). Más tarde me ocuparé de esta categoría con especial atención, y trataré de justificar lo que ahora voy a aceptar como un axioma que más adelante matizaré convenientemente: un trastorno de la personalidad suele ser resultado de una educación que arruina las potencialidades de una persona para participar normalmente en nuestra sociedad, al resultado de una infancia que llega a convertir a una persona en alguien incapacitado para actuar mediante juicios sobre valores convencionalmente admitidos en nuestra comunidad moral. Por último, considero *menor de edad* a todo aquel que no ha terminado el proceso de educación que debería convertirle en persona consciente de lo que es justo e injusto, bueno y malo, y capaz de medir razonablemente el alcance de sus acciones.

A modo de pequeño apéndice para este apartado me gustaría recordar que a veces usamos un tipo de excusas que no son válidas que constituyen un género por sí mismas.⁶⁰⁸ Se trata de expresiones tales como “sí pero es tan gracioso”, “sí pero es tan maja”, “sí pero es tan guapo”; o también “ya le conoces”, “ella es así”, “a su edad ya no va a cambiar”. Aunque sin duda el nombre que más le conviene es el de “excusas estúpidas”, como se recordará bauticé a este grupo como el de las *disculpas*. No lo he incluido en el gráfico porque, aunque funciona en la vida diaria, no constituye una forma legítima de

⁶⁰⁸ Ya hablé de ellas en las páginas 287 y 288.

exculpación, y sobre todo porque, aún siendo una variación bastarda de las excusas, forma parte de la actitud reactiva y es una forma de exculpación *emotiva* de responsabilidad. Insisto en que las disculpas no son formas de exculpación válidas, pero se dan, y por ello he querido dejar de nuevo constancia de ellas aquí.

En cualquier caso puede rastrearse el alcance de las categorías que he señalado en los ejemplos de expresiones que pone Strawson en “Freedom and Resentment”, en este aspecto le he sido fiel y no he hecho sino clasificar más claramente sus intuiciones (si se obvia el hecho de que he inventado el género “disculpas”). Y con esta tarea voy a continuar en el siguiente apartado, intentando sacar por fin a la luz en qué consiste más exactamente lo que Strawson llama de forma harto imprecisa una “comunidad moral”.

III-4. Vivir en una comunidad moral

En otras palabras, los androides con la nueva unidad cerebral Nexus-6 —desde un punto de vista pragmático y nada disparatado— sobrepasaban a una considerable porción de la humanidad, aunque fueran los del nivel inferior. [...] Pero había nuevos criterios, por ejemplo el test de empatía de Voigt-Kampff. Un androide, por dotado que estuviera en cuanto a capacidad intelectual pura, no podía encontrar el menor sentido en la fusión que experimentaban rutinariamente los [seres humanos] (...).⁶⁰⁹

Un grupo de psiquiatras de Leningrado pidió a la WPO que aplicara el método de perfil de personalidad más moderno y preciso para determinar la presencia de un androide, o sea la escala de Voigt-Kampff, a un grupo seleccionado de pacientes humanos, esquizoides y esquizofrénicos. Especialmente aquellos que revelan lo que se denomina un «achataamiento del afecto». [...] Los psiquiatras de Leningrado (...) creen que una pequeña proporción de seres humanos no podría pasar la prueba de Voigt-Kampff. Si los sometiera usted al test en el curso de una tarea policial, quedarían clasificados como robots humanoides.⁶¹⁰

⁶⁰⁹ Dick, *Sueñan los androides con ovejas eléctricas*, 1992: 12.

⁶¹⁰ Dick, *Sueñan los androides con ovejas eléctricas*, 1992: 37-38.

En el apartado anterior se ha visto que cabe que alguien actúe voluntariamente y sin embargo no sea responsable de su acción. Esto es así porque no puede entenderse la responsabilidad de forma unívoca, sino que es necesario comprender que hay diferentes formas de ser responsable de algo. Una de ellas está directamente unida a la posibilidad de actuar libremente en el sentido de actuar en ausencia de factores externos que limiten nuestra capacidad de elección. La otra va ligada a la posibilidad de actuar libremente en el sentido de actuar dándonos la posibilidad de elegir racionalmente, de acuerdo con las expectativas de aquellos con quienes convivimos. De estos dos aspectos de la responsabilidad, que permiten comprender mejor qué es ser responsable y qué es vivir en una comunidad moral,⁶¹¹ voy a ocuparme en esta parte, y con ello daré por terminado mi estudio del lugar en que es posible la responsabilidad moral: el mundo participante.

III.4.1. La personalidad y su virtud: las dos caras de la responsabilidad

En II.1.4. me ocupé de la visión utilitarista de la responsabilidad moral y decidí dejarla de lado porque no tiene en cuenta aspectos esenciales de dicha responsabilidad. Consideré que el utilitarismo no da cuenta de aquello en que consiste realmente hacer atribuciones de responsabilidad, pues entiende este hecho como un mero mecanismo para el progreso social. Ahora debo en parte reivindicar que las atribuciones de responsabilidad son *también* mecanismos

⁶¹¹ Voy a usar en estas páginas con total ligereza el concepto de comunidad, creo que mi uso del mismo queda totalmente fuera de la clásica disputa entre liberales y comunitaristas (hasta el punto de que ha menudo utilizaré, como vengo haciendo hasta ahora, las nociones de “teoría de la justicia” y “teoría de la virtud” como fundamentalmente sinónimas o cuando menos complementarias). No obstante, no puedo ignorar los progresos que, acerca del concepto de comunidad, han hecho muchos pensadores y por ello me permito recomendar aquí la lectura de algunas obras que se ocupan de ello, a saber: MacIntyre, A. 1981, *After Virtue*, London: Duckworth; Sandel, M. 1982, *Liberalism and the Limits of the Justice*, Cambridge: Cambridge University Press; Taylor, Ch. 1990, *Sources of the Self*, Cambridge: Cambridge University Press; Walzer, M., 1983, *Spheres of Justice*, New York: Basic Books.

para corregir a los individuos implicados en una determinada acción que juzgamos moralmente, así como, mediante la disuasión por un castigo ejemplar o cosas similares, mecanismos para la educación de la sociedad. Sin duda el sistema de excusas y exenciones fruto de un desarrollo moral basado en las emociones reactivas necesita de ese carácter regulador social de la responsabilidad. Ahora bien, no se debería dejar de tener en cuenta que hacer atribuciones de responsabilidad, como intenté mostrar en *II.1.4.*, es algo más que un mecanismo de regulación social, y constituye en sí mismo un juicio concreto formulado sobre una persona concreta y por una acción concreta comprendida como evaluable por sí misma, y no sólo en relación a sus consecuencias sociales más o menos halagüeñas. No es cierto que nuestras emociones reactivas consistan en expresiones de un anhelo de una sociedad más justa, aunque sí es cierto que son fruto de ese anhelo que tenemos interiorizado y que es la base de las expectativas en que esas emociones se asientan. Sin duda reformamos al delincuente con un ojo utilitarista, pero le juzgamos con algo más que ese ojo. La cárcel no es sólo un reformatorio, sino un lugar de castigo.

Bien, esta doble cara del juicio moral es la expresión de la doble cara que tiene el concepto de responsabilidad. Por un lado existe un tipo de responsabilidad que consiste en la capacidad para llevar a cabo una acción voluntaria, libremente y con conocimiento de causa. Por otro lado existe un tipo de responsabilidad que consiste en la capacidad para interpretar el entramado social de expectativas y actuar de acuerdo con él. Una misma persona puede ser responsable por una acción desde estos dos puntos de vista, pero también desde uno sólo de ellos; pero para que un agente sea juzgado plenamente responsable por una acción debe serlo desde ambos. Sin ir más lejos las excusas dicen que el agente no obró libremente o con conocimiento de causa, pero no que el agente sea incapaz de jugar el juego social de las expectativas que componen una comunidad moral. Y en lo que respecta a las

exenciones ocurre lo contrario: el agente actúa libremente, según sus deseos más profundos, pero no es capaz de comprender el juego social de expectativas en el que está inmerso.

A partir de aquí voy a considerar que *ser miembro de una comunidad moral es ser capaz de comprender el juego social de expectativas al que jugamos continuamente en nuestras relaciones interpersonales ordinarias*. Ese juego entraña la asunción de ciertos valores básicos que conforman una teoría de la justicia o de la virtud implícita, convencionalmente aceptada y mediante la cual somos capaces de distinguir el bien del mal, lo justo de lo injusto. Ser miembro de una comunidad moral es ser capaz de comprender que hay que intentar, en la medida de lo posible, cumplir una expectativa de justicia o de bondad que todo aquel que sea miembro de la comunidad moral tiene, esa expectativa de buena voluntad de la que habla Strawson. Ser miembro de una comunidad moral implica, pues, ser capaz de tener así mismo expectativas puestas en los demás sobre cuya violación uno podría pedir cuentas porque uno mismo sería capaz de rendirlas si los demás se las solicitasen. El enfermo mental, aquel que sufre un trastorno de personalidad y el niño no forman parte de esa comunidad moral porque ni son capaces de dar cuenta de sus propias expectativas, ni son capaces de comprender las de los demás. En cierto sentido no conocen el código de la circulación social. Actúan libremente, pero con una libertad insignificante, puramente agencial, que no es suficiente para que sean considerados responsables moralmente.

En I.3.2. me ocupé de la personalidad, concretamente del carácter, ese lugar donde viven nuestros, por decirlo a la Frankfurt, deseos de segundo orden. Ahora, por fin, me corresponde cerrar la investigación abierta entonces. El que una acción sea fruto de nuestro carácter, de nuestros deseos de orden superior, no es más que la mitad de los requisitos necesarios para la responsabilidad moral. La otra mitad es que esa acción encaje con una teoría de la virtud, con un sistema de valores, esto es, que sea fruto de un conjunto

organizado y funcional de principios que nos permita relacionarnos de manera correcta con los miembros de la comunidad moral de que formemos parte. Así, puede decirse que *la personalidad, unidad del agente y sujeto de la libertad, se compone de carácter y de virtud*. El primero es sobre todo la base de la libertad de la voluntad, la segunda es indispensable para la responsabilidad moral.

Siguiendo la idea de Frankfurt de que existen deseos de primer orden y deseos de segundo orden, Gary Watson propone que distinguir entre un sistema de valores y un sistema de motivos, siendo ambos sistemas requisitos necesarios para que un agente pueda ser considerado responsable:

El sistema de valores de un agente es ese conjunto de consideraciones que, combinadas con sus creencias (y estimaciones de probabilidad), produce juicios de esta forma: aquello que yo debería hacer en estas circunstancias, teniendo todo en cuenta, es *a*. Adscribir libre albedrío a un ser presupone que se trata de un ser capaz de hacer juicios de este tipo. Para ser ese tipo de ser, uno debe asignar valores a estados de cosas alternativos, esto es, jerarquizarlos en términos de valor.⁶¹²

El sistema de motivos de un agente es ese conjunto de consideraciones que le mueven a actuar.⁶¹³

Tal y como estoy tratando la cuestión puede decirse que *el sistema de motivos es el carácter del agente y el sistema de valores es su teoría de la virtud o de la justicia*. Así, *la personalidad se compone de un sistema de motivos y de un sistema de valores*. Según Gary Watson un agente no es responsable cuando su sistema de valores y su sistema de motivos no coinciden. Como se ha visto al tratar de las excusas y de las exenciones, esto puede suceder de varias formas: o bien cuando alguien se ve obligado a hacer algo que no encaja con su sistema de valores por fuerza mayor o por ignorancia (y entonces le excusamos); o bien cuando alguien directamente carece de un sistema de valores (y entonces le eximimos). Decir de alguien que carece de un sistema de valores es otra forma más de decir que está fuera de la comunidad moral, pues la expectativas que articulan nuestra vida participante y por lo tanto nuestra comunidad moral, se crean y se justifican gracias a que cada uno tiene un cierto sistema de valores.

⁶¹² Watson 1975, en Watson 1982: 105.

Así pues, la responsabilidad tiene dos caras que se basan en dos sistemas de deliberación o, siguiendo a Frankfurt, dos niveles de volición, que son los que nos permiten funcionar normalmente en una comunidad moral: “(...) el estatus de una persona como agente responsable no requiere solamente la capacidad para conformar sus deseos y conductas a sus valores profundos (...) sino también la capacidad para adquirir los valores *correctos* —esto es, aquellos por los que le consideramos responsable de tener.”⁶¹⁴

Voy ahora, una vez comprendido por fin en qué consiste la intuición strawsoniana de una comunidad moral, a estudiar esas dos caras de la responsabilidad.

La responsabilidad personal

Varios autores (Harry Frankfurt, Gary Watson, Charles Taylor) han insistido en cierto aspecto de la responsabilidad que es el que me va a ocupar aquí. Susan Wolf los ha bautizado como los defensores del “punto de vista del yo profundo” o “del verdadero yo”. Estos autores se han esforzado por investigar aquello en que consiste ser una persona con libre voluntad y por ello aquí, a ese sujeto que es una persona libre, voy a considerar que es al que se le puede atribuir un cierto tipo de responsabilidad: la responsabilidad personal. Basándome en las que he considerado excusas tipo, ignorancia y fuerza mayor, voy a definir esa responsabilidad personal. *Es responsable personalmente de sus actos todo aquél que actúa por propia voluntad con conocimiento de causa y sin ser coaccionado.*

Se dice de alguien que *actúa con conocimiento de causa* cuando es plenamente consciente de las consecuencias de su acción, y por lo tanto la lleva a cabo porque esas consecuencias son el estado de cosas que él quiere que se dé debido a su acción. *No* es responsable si, de haber sabido lo que iba a desencadenar su acción, no la hubiese llevado a cabo. Actuar con

⁶¹³ Ibidem: 105-106.

⁶¹⁴ Watson 1996: 228.

conocimiento de causa supone no ver mermadas por ignorancia mis posibilidades de acción.

Se dice de alguien que actúa *sin ser coaccionado* cuando no actúa forzado por algo ajeno a su voluntad. Cuando, de no haberse dado ciertas circunstancias tales como una amenaza o un grave riesgo para él, habría elegido actuar de modo diferente a como lo hizo efectivamente.

Retomando la distinción de Gary Watson entre un sistema de valores y un sistema de motivos, se puede decir que alguien es responsable personalmente cuando actúa de acuerdo a ambos sistemas, esto es, cuando nada le impide que su deseo concuerde con sus motivos y sus valores. Alguien no es responsable personalmente de su acción cuando la lleva a cabo porque no tiene más remedio y no porque realmente lo desee, porque encaje con sus aspiraciones y deseos más profundos.

Este es el tipo de libertad básico que supone el actuar voluntariamente. Todo agente debe ser capaz de reconocerse en su acción porque toda acción, para ser libre en el sentido que estoy tratando, debe ser fruto de algo así como su “verdadero yo”. Cuando actuamos y más tarde somos capaces de justificar racionalmente que no queríamos haber obrado del modo en que lo hicimos pero que “no tuvimos más remedio” o que “no pudimos evitarlo”, entonces no puede decirse que seamos plenamente responsables de nuestra acción. Seremos sin duda causalmente responsables, pero no *moralmente* responsables porque no habremos podido actuar de acuerdo con nuestro carácter, desde nuestra “responsabilidad personal”.

La responsabilidad areteica⁶¹⁵

Aunque este apartado pretende precisamente llegar a explicar detalladamente en qué consiste lo que llamo “la responsabilidad areteica”,

⁶¹⁵ Tomo esta denominación de Gary Watson 1996.

quiero suministrar aquí una primera aproximación a dicho concepto a fin de que el lector pueda atribuirle algún significado, aunque sea de forma provisional. Hay que entender la responsabilidad areteica como aquella que hace referencia a nuestra facultad para actuar dentro del orden, de las reglas del juego, del sistema de valores que configura nuestra vida participante. Esto puede entenderse como ser capaz de participar de una teoría de la virtud (*aretê*), esto es, como vivir dentro de una estructura social que promueve ciertos patrones de comportamiento que considera correctos y útiles para el florecimiento de la comunidad hasta convertirlos en hábito. Así, no está directamente ligada, como sí lo está la responsabilidad personal, a mi carácter, sino a la forma en que éste encaja con el mundo.

Más arriba, en *III.3.2.*, defendí que aquel a quien aplicamos una exención de responsabilidad no ve mermada su voluntad. En efecto, el enfermo mental, el trastornado o el niño no actúan viendo menoscabado su sistema de motivos, el problema reside en que ese sistema de motivos funciona sin un sistema de valores que sea capaz de regular los motivos según las normas básicas de una comunidad moral. Eso, siguiendo a Susan Wolf, constituye una forma de “enajenación”:

[E]n orden a ser responsable, un agente debe estar *cuerdo*. Ordinariamente no está en nuestro poder determinar si estamos o no cuerdos. Muchos de nosotros, parece, somos afortunados, pero algunos de nosotros no lo son. Es más, estar cuerdo no quiere decir necesariamente que uno tenga algún tipo de poder de control del que una persona enferma carece. Para nuestra desgracia, alguna gente enferma, (...), puede tener un control completo sobre sus actos, e incluso un control completo sobre la personalidad que le lleva actuar. El deseo de estar cuerdo no es entonces un deseo por otra forma de control. Es más bien un deseo de que la propia personalidad esté conectada con el mundo de cierta manera –podríamos incluso decir que es un deseo de que la propia personalidad esté *controlada por* el mundo de ciertas maneras y no de otras.⁶¹⁶

En definitiva, la persona enajenada aspira a tener un sistema de valores que le permita tener las mismas referencias con las que cuentan aquellos con quien vive. Sin ellas es imposible participar del juego de expectativas en que

⁶¹⁶ Wolf 1988, en Kane 2002: 154.

todos estamos, pues ese juego de expectativas requiere algo más que la capacidad para actuar atendiendo a motivos, requiere también el conocimiento de un código que nos dice qué vale y qué no vale a la hora de actuar, y sobre todo qué vale y qué no vale a la hora de enjuiciar una acción. No hay, pues, únicamente un yo profundo que tiene deseos de segundo orden, sino también un yo “cuerdo”, esto es: un yo integrado en la sociedad, que sabe lo que está bien y lo que está mal y que no es resultado de una infancia atormentada o de una sociedad pervertida. Así, es necesario pertenecer a cierto mundo, a cierta comunidad moral que nos dote de valores y deseos que hagan que funcione la responsabilidad. “Consecuentemente, para que un agente tenga un motivo particular que esté sujeto a cualquier evaluación moral, el agente debe poseer un dominio de las interpretaciones convencionales de las acciones-tipo que expresan la comprensión de una comunidad moral de las prohibiciones y expectativas de la moralidad. Sólo en ese contexto puede un agente *moral* actuar desde motivos que están sujetos a evaluación moral.”⁶¹⁷ Cuando no puede actuar dentro de ese contexto puede decirse de esa persona que está enajenada en el sentido etimológico de la palabra, y así es como cabe leer el no pertenecer a una comunidad moral: como ser extranjero en un mundo regido por un determinado sistema de valores, como tener por ajenos e incomprensibles los principios de las “relaciones interpersonales normales” de la sociedad en que se vive, que es vista moralmente como un país extraño.

Susan Wolf explica que para que alguien sea considerado cuerdo, esto es, un miembro de una comunidad moral, debe, al actuar, saber qué es lo que hace, pero también poder esperar cabalmente que sus valores encajen con los del resto del mundo y que por lo tanto sus acciones no tienen un significado diferente para el resto del mundo del que podrían tener para él. “Poniendo estas dos condiciones juntas, debemos comprender la cordura, entonces, como

⁶¹⁷ McKenna 1998b: 5.

la habilidad mínima suficiente para, cognitiva y normativamente, reconocer y apreciar el mundo por lo que es.”⁶¹⁸ Quien no cuenta con esta habilidad no está cuerdo, quien no está cuerdo no forma parte de una comunidad moral, y quien no forma parte de una comunidad moral no puede ser considerado moralmente responsable sino tan sólo causalmente responsable. Esa persona no es capaz de *ser areteicamente responsable, esto es, no es capaz de actuar según una teoría de la virtud (o de la justicia), un sistema de valores comúnmente aceptado*.⁶¹⁹

Con esto se podría dar por terminado el cuadro de la responsabilidad areteica, sin embargo no me gustaría dejar de hacer aquí una observación no exenta de importancia. El hecho de que haya hablado de cordura no implica que sólo el enfermo esté fuera de la comunidad moral, sobre todo si se entiende “enfermo” en sentido estricto vinculándolo a la idea de patología. También el niño, que todavía no ha conseguido estructurar suficientemente, por su minoría de edad, su sistema de valores, está fuera de la comunidad moral. Y fuera de la comunidad moral se encuentra también el que llamé más arriba “persona con un trastorno de la personalidad”, refiriéndome no sólo a personas con trastornos tipificados, sino en general también a todas aquellas personas que, por el ambiente en que crecen, se ven incapaces de llegar a formar parte de la comunidad moral. La educación debería suministrar una cierta competencia para distinguir entre el bien y el mal dentro de un código compartido por la mayor parte de nosotros. El niño no ha recibido aún esa educación, o no ha practicado lo suficiente como para que se le considere plenamente miembro de la comunidad moral. Algunas personas han completado su educación, pero ésta fue deficiente y no les dotó de un sistema de valores manejable. De estos últimos, enajenados morales, no puede decirse

⁶¹⁸ Wolf 1988, en Kane 2002: 155.

⁶¹⁹ Poseer ese sistema de valores (teoría de la virtud, justicia...) no implica que se pueda hacer explícito, ni siquiera que se sea plenamente consciente de él, ni que se tenga perfectamente jerarquizado. Se trata de hábitos, de costumbres morales, de patrones básicos de comportamiento, más bien asumidos que aprendidos.

que estén realmente cuerdos en el sentido visto más arriba de la mano de Susan Wolf. En cualquier caso, como se ve aparecen aquí las variedades de persona que, según el análisis que hice más arriba, componen mi catálogo de exenciones; con lo que se puede decir que, en definitiva, estar cuerdo es la medida para ser considerado areteicamente responsable y la enajenación, en sus diferentes variantes y entendida en un sentido muy amplio, es la condición necesaria para que alguien sea eximido de responsabilidad moral.

III.4.2. Los que viven fuera de la comunidad moral

Tras tantas y tantas páginas probablemente el lector ya no lo tendrá en su mente, sin embargo convendría que hiciese aquí el esfuerzo de recordar en qué sentido hablé de “fenomenología” en la presentación del trabajo. Resulta cuando menos sugerente el hecho de que algunos criminólogos hayan optado también por usar el método fenomenológico para esclarecer los motivos de algunos asesinos y el significado de sus crímenes: “[...] -la fenomenología empírica- se centra en las experiencias individuales porque, al interpretar la forma en que un individuo relata ciertas situaciones, podemos determinar cuáles son las estructuras subyacentes a esas experiencias. Las narraciones personales nos revelan los significados que organizan la vida de un individuo.”⁶²⁰ Es realmente alentador ver que algunos científicos, a la hora de estudiar la personalidad, proceden de manera similar a como se ha procedido a lo largo de este trabajo: centrándose en las experiencias morales de los individuos, intentando descubrir sus “estructuras subyacentes” y el entramado de significados con que construyen su vida. La filosofía no debería dar la espalda a las ciencias naturales, porque tanto la una como las otras aspiran a lo mismo: a conocer mejor el mundo. Así, si los filósofos realmente queremos

aprender algo acerca de los sujetos y sus emociones, sus deseos y sus creencias, conviene que tengamos en cuenta lo que psiquiatras y psicólogos pueden enseñarnos a este respecto. Por ello aquí voy a atender a resultados recientes de la psiquiatría y la criminología en lo que concierne al estudio de los trastornos de la personalidad, porque “[l]o mismo que los puntos extremos de un continuo ayudan a definir el centro, el conocimiento de (...) forma[s] extrema[s] del comportamiento violento mejora nuestra comprensión de la condición humana.”⁶²¹ En este trabajo se han tocado algunos aspectos de la condición humana y me gustaría ahora estudiar sus límites para ver hasta qué punto lo que la fenomenología filosófica ha ido desvelando no es más que un constructo teórico que no encaja con la realidad, o por el contrario es capaz de dar cuenta de algunos datos experimentales. Afortunadamente para la teoría que ha sido desarrollada en este trabajo, muchas de las conclusiones a las que importantes investigadores en psiquiatría y criminología han llegado en sus estudios sobre personas que “viven fuera de una comunidad moral”, encajan considerablemente bien con las conclusiones de la teoría aquí desarrollada. Así, toca ahora estudiar lo que supone ser miembro de una comunidad moral desde el punto de vista de la psiquiatría. Me voy a centrar en los estudios realizados con psicópatas, pues la psicopatía puede tomarse como un caso paradigmático de trastorno de la personalidad, esto es, en los términos de este trabajo, como un perfecto ejemplo de aquello en que consiste vivir fuera de una comunidad moral.

El cerebro del psicópata

La escala Hare (PCL-R – *Psychopathy Checklist-Revised*) se ha convertido en el mejor elemento para diagnosticar la psicopatía, y por lo tanto también para definirla. Dicha escala consta de veinte aspectos que son rasgos típicos de la

⁶²⁰ Scrapeck, C. A. “Los motivos del asesino en serie”, en Raine & Sanmartín 2000: 167.

⁶²¹ Ibidem: 158.

personalidad de la mayor parte de los psicópatas. Los problemas que conlleva la psicopatía se dan en tres ámbitos interdependientes: el de las relaciones sociales, el de la afectividad y el de la conducta. “En el primer ámbito, los psicópatas tienden a manipular y engañar a los demás. En su vertiente afectiva, adolecen de empatía: son incapaces de ponerse en lugar del otro. Finalmente, su comportamiento es antisocial.”⁶²² Durante mucho tiempo este último aspecto de la psicopatía fue el más destacado, lo cual dificultaba la tarea de distinguir psicopatía y trastorno antisocial de la personalidad. Sin embargo el comportamiento antisocial no es tanto un rasgo definitorio de la psicopatía como una consecuencia prácticamente ineludible de la misma: “Los principales inhibidores de la violencia y la conducta antisocial (empatía, vínculos emocionales, miedo al castigo, sentimientos de culpa, etc.) son inexistentes o muy deficientes en los psicópatas. En cambio, el egocentrismo, el narcisismo, la autojustificación, la impulsividad, la falta general de inhibiciones comportamentales y la necesidad de poder y control, constituyen lo que se podría llamar la fórmula perfecta para los actos antisociales y criminales.”⁶²³ El extremo más conocido de la psicopatía lo constituyen los asesinos en serie, una figura terrorífica de la que el cine viene abusando de tal modo en los últimos años que ha llegado a convertirse en el ejemplo paradigmático de trastorno de la personalidad. Sin embargo los psicópatas representan tan sólo el 1% de la población general (el 25% de la población reclusa), y una mísera parte de ese uno por ciento lo constituyen los asesinos en serie.

Existe una serie de evidencias que permiten entender la psicopatía desde el punto de vista fisiológico y que arrojan cierta luz sobre el comportamiento violento de los psicópatas. En lo que respecta a las características fisiológicas de la psicopatía, experimentos con tomografías de

⁶²² Raine & Sanmartín 2000: 8.

⁶²³ Hare, R. D. “La naturaleza del psicópata: algunas observaciones para entender la violencia depredadora humana”, en Raine & Sanmartín 2000: 24.

emisión de positrones revelan que los psicópatas presentan, en comparación con personas normales, una baja actividad de la corteza prefrontal⁶²⁴, lo cual es un factor de predisposición a la violencia porque, entre otras razones, “un funcionamiento prefrontal reducido puede traducirse en una pérdida de la inhibición o control de estructuras subcorticales, filogenéticamente más primitivas, como la amígdala, que se piensa que está en la base de los comportamientos agresivos”; y también, “la pérdida de flexibilidad intelectual y de las habilidades para resolver problemas, así como la merma de capacidad para usar la información suministrada por indicaciones verbales que nacen del mal funcionamiento prefrontal, pueden deteriorar seriamente habilidades sociales necesarias para plantear soluciones no agresivas a los conflictos”.⁶²⁵

Nuevos experimentos con una muestra mayor de psicópatas⁶²⁶, además de corroborar los ya conocidos, aportaron nuevos datos sobre el funcionamiento de sus cerebros. Así, se descubrieron problemas en el giro angular izquierdo, lo cual se asocia a “problemas en la capacidad verbal” y “deficiencias de lectura y cálculo”⁶²⁷, ambas cosas pueden acarrear dificultades de aprendizaje y con ello fracaso escolar y profesional que pueden predisponer a la violencia. También se encontró que la actividad del cuerpo calloso era menor entre los asesinos. A este respecto la hipótesis es que el cuerpo calloso, un conjunto de fibras nerviosas que une los dos hemisferios cerebrales, es el responsable de que el hemisferio izquierdo no sea capaz de regular adecuadamente las emociones negativas, de cuya generación se supone responsable al hemisferio derecho: “Es interesante constatar a este respecto que el hemisferio derecho es el dominante en las ratas que padecen estrés en

⁶²⁴ Raine, A.; Buchsbaum, M. S.; Stanley, J.; Lottenberg, S.; Abel, L. y Stoddard, J. 1994. “Selective reductions in pre-frontal glucose metabolism in murderers”, *Biological Psychiatry* 36: 365-373.

⁶²⁵ Raine, A. “Psicopatía, violencia y neuroimagen”, en Raine & Sanmartín (eds.) 2000: 65.

⁶²⁶ Raine, A.; Buchsbaum, M. S. & La Casse, L. 1997. “Brain abnormalities in murderers indicated by positron emission tomography”, *Biological Psychiatry* 42: 495-508.

⁶²⁷ Raine, A. “Psicopatía, violencia y neuroimagen”, en Raine & Sanmartín 2000: 66.

las primeras fases de su existencia y que estas ratas incurren en tasas altas de muricidio (...) tasas [que] aumentan en el caso de las ratas a las que se les ha seccionado el cuerpo calloso (...).”⁶²⁸ Por último, la amígdala, el hipocampo y el tálamo de los asesinos analizados demostraron funcionar de manera anormal. “La amígdala, el hipocampo y la corteza prefrontal se integran en el sistema límbico que gobierna la expresión de las emociones, a la vez que el tálamo transmite *inputs* desde las estructuras subcorticales límbicas hasta la corteza prefrontal (...).”⁶²⁹ Además estas estructuras intervienen en el aprendizaje, la memoria y la atención, y el hecho de que estén dañadas puede conducir a que los asesinos psicópatas sean incapaces de aprender de la experiencia, lo cual explicaría la altísima tasa de reincidencia de estos delincuentes. “La amígdala juega además un papel importante en el reconocimiento de los estímulos afectivos y socialmente significativos, por lo que su destrucción se traduce en una carencia de miedo (...).”⁶³⁰

Una de las preguntas interesantes que puede surgir a la luz de estos datos es la siguiente: el daño que registran esas regiones del cerebro en los psicópatas, ¿es congénito o fruto de factores ambientales? La mayor parte de los investigadores consideran que la psicopatía sólo puede explicarse por la concurrencia de ambas dimensiones:

Si las deficiencias cerebrales son las que contribuyen directamente a la aparición de la violencia, ¿cuáles son, a su vez, sus causas? Por una parte [...] quizá haya eventos (...) que se produzcan en las fases iniciales de la vida de los asesinos y que contribuyan a la aparición de disfunciones cerebrales. Por ejemplo, hay investigaciones que ponen de manifiesto que los delincuentes violentos suelen provenir de hogares en los que hay maltrato. [...] Por otra parte [...] la genética del comportamiento y la genética molecular están haciendo grandes contribuciones a la clarificación del crimen y de la violencia.⁶³¹

En cualquier caso hay algo claro en lo que apenas hay voces discordantes: sólo puede hablarse de *predisposición* genética a la violencia, pero

⁶²⁸ Ibidem: 67.

⁶²⁹ Ibidem: 68.

⁶³⁰ Ibidem.

no de *predeterminación*. El que alguien predispuesto a la violencia por ciertos rasgos fisiológicos de nacimiento, acabe siendo efectivamente una persona violenta depende de que haya un caldo de cultivo social en que esa predisposición se actualice, y esto es lo que interesa aquí. El terreno de las predisposiciones genéticas cae fuera del alcance del filósofo, en cambio, la forma en que los individuos adquieren un sistema de valores, incluso aquellos que tienen defectos congénitos que les dificultan su adquisición, está directamente ligada a la adquisición del estatus de persona moralmente responsable. Y en este punto es donde es relevante el uso de la filosofía.

Las emociones del psicópata

Como se ha ido viendo a lo largo de este trabajo, ser moralmente responsable viene a coincidir con ser capaz de vivir en el mundo participante. Ello implica que uno es capaz de tener expectativas y atribuírselas a los demás, y de juzgar el cumplimiento de las expectativas propias y ajenas. Como se ha visto este juicio no es meramente racional en el sentido clásico del término, sino que está transido por emociones, de tal forma que a veces, en el día a día casi siempre, esas emociones constituyen en sí mismas juicios morales.

Más arriba hablé del hecho de que todos los que formamos parte de una comunidad moral poseemos un sistema de valores. Dicho sistema ha de ser entendido no como un código moral concreto, sino como una facultad: la facultad que nos permite sentir las emociones adecuadas en los momentos apropiados; o en general emitir los juicios adecuados, emotivos o no, en los momentos apropiados. Parafraseando a Susan Wolf: se trata de que nuestro mundo moral coincida suficientemente con el mundo moral de quienes nos rodean como para que seamos capaces de interpretar de forma similar experiencias similares, y por lo tanto hacer juicios de valor semejantes. Para

⁶³¹ Ibidem: 70-71.

evitar las connotaciones comunitaristas (y el peligro de relativismo consiguiente) que podría tener la idea de una comunidad moral, se puede señalar el siguiente hecho: “Cabe destacar que la relación entre psicopatía y violencia no es sólo propia de la psiquiatría occidental, sino que se da en todas las culturas. En todas ellas, por tanto, el potencial para la violencia es un síntoma de trastorno (...).”⁶³² Tal vez existan, como señala Allan Gibard en *Wise Choices, Apt Feelings*, comunidades morales basadas en la vergüenza y no en la responsabilidad, pero en todos los lugares del mundo se marcan diferencias entre gente que “funciona” correctamente desde el punto de vista moral, y gente que “no funciona” moralmente. Esto radica en el hecho de que, algunas personas, no es que tengan sistemas de valores alternativos al que propone la comunidad moral, sino que carecen de la capacidad para tener un sistema de valores. El rebelde es alguien que conoce y comprende un sistema de valores, que es capaz de pensar un mundo posible regido por un sistema de valores diferente, y que lucha por sustituir el mundo real por ese mundo posible. Para el psicópata, para el que está fuera de la comunidad moral, no hay mundos posibles alternativos, tan sólo existe un mundo y no es suyo. El rebelde pertenece a la comunidad moral, desde ella opera y se rebela. El psicópata no pertenece a la comunidad moral, vive fuera de ella. Él cree que se rebela, que cuando asesina está haciendo justicia, pero no es así. También cree que no hecho nada malo al asesinar a alguien, y si reconoce haberlo hecho no es porque entienda aquello en que consiste hacer algo malo, sino porque es capaz de reconocer un hecho por el que, según la ley, es susceptible de ser castigado. “Algunas personas creen a los psicópatas cuando dicen tener remordimientos, pero los observadores más astutos no confían en sus palabras

⁶³² Hare, R. D. “La naturaleza del psicópata: algunas observaciones para entender la violencia depredadora humana”, en Raine & Sanmartín 2000: 18, nota 1. Entre los estudios que confirman esta afirmación conviene consultar Murphy, J. M. 1996. “Psychiatric labelling in cross-cultural perspective: Similar kinds of disturbed behavior appear to be labelled abnormal in diverse cultures”, *Science* 191: 1019-1028.

porque éstas parecen ser una simple imitación de algo que, en realidad, no entienden.”⁶³³ Cuando un psicópata habla de remordimiento está actuando, está adaptando su comportamiento lingüístico a las circunstancias. Al igual que John Searle dentro de la habitación china, el psicópata que habla de sentimientos morales está hablando chino sin entenderlo, opera a nivel sintáctico pero no semántico: “(...) un asesino convicto escribió: «Hay emociones –todo un amplio abanico de ellas- que sólo conozco a través de las palabras, de la lectura y de mi imaginación inmadura. Puedo imaginarme que siento estas emociones, pero no las siento».”⁶³⁴ Ese asesino fue puesto en libertad y volvió a asesinar. Esto dijo acerca de su víctima, un camarero aspirante a actor: “No hubo dolor, fue una herida limpia... No tenía futuro como actor. Lo más probable es que hubiera acabado en otro tipo de trabajo.”⁶³⁵ Se ve claramente que la mirada al mundo de este sujeto es amoral: “[n]os encontramos frente a un hombre, por lo demás inteligente, que describe un asesinato impulsivo desapasionada y casi flemáticamente. No logra comprender a qué viene tanto lío.”⁶³⁶ Los seres humanos moralmente normales y los psicópatas no compartimos el mismo mundo moral. Sin duda el uso y abuso que viene haciéndose del concepto de inconmensurabilidad podría levantar ciertas sospechas sobre lo que estoy diciendo, no obstante hay experimentos que corroboran este discurso: los psicópatas no entienden muchos juicios morales porque son incapaces de descodificar la carga emotiva que poseen muchas palabras. Este punto ha sido corroborado experimentalmente mediante diversas técnicas: el registro de *event-related potentials* (ERP – potenciales de sucesos relacionados), la tomografía

⁶³³ Hare, R. D., “La naturaleza del psicópata: algunas observaciones para entender la violencia depredadora humana”, en Raine & Sanmartín 2000: 41.

⁶³⁴ Ibidem.

⁶³⁵ Ibidem.

⁶³⁶ Ibidem.

computarizada de emisión de fotones (SPECT), y la resonancia magnética funcional (RMf).

“Williamson y otros⁶³⁷ registraron el tiempo de reacción y los ERP – breves cambios en la actividad eléctrica del cerebro asociados con el procesamiento cognitivo y la evolución de un estímulo, en este caso, palabras de una tarea de decisiones léxicas en la que los delincuentes miraban una pantalla de ordenador donde aparecían y desaparecían cadenas de letras.”⁶³⁸ En el experimento participaron delincuentes psicópatas y no psicópatas, y un grupo de control de no criminales, todos debían pulsar un botón, tan rápida y certeramente como pudieran, a la vista de cadenas de letras que formaban una palabra existente. Esas cadenas de letras podían formar grupos de sílabas impronunciables, palabras inexistentes pero pronunciables y por último palabras con sentido. Dentro de éstas las había con contenido neutro y con contenido emocional (tanto positivo como negativo). Tanto los tiempos de reacción como los ERP fueron muy diferentes entre los delincuentes no psicópatas y el grupo control, por un lado, y los psicópatas, por otro. A diferencia de lo que ocurría en el caso de los primeros, que reaccionaban más rápidamente ante palabras con contenido emocional que con contenido neutro, “en los psicópatas no había ninguna diferencia ni en el tiempo de reacción ni en los ERP ante las palabras neutras y con contenido emocional”.⁶³⁹ La no diferencia en los ERP de las regiones centrales y parietales del cerebro de los psicópatas tenía lugar sobre todo en los componentes tardíos, pequeños y breves, “quizá porque procesaban la información rápida y superficialmente, y porque hacían poco más que tomar una decisión léxica, mientras que los no psicópatas seguían procesando o

⁶³⁷ Williamson, S. E.; Harpur, T. J. & Hare, R. D. 1991. “Abnormal processing of affective words by psychopaths”, *Psychophysiology* 28: 260-273.

⁶³⁸ Hare, R. D. “La naturaleza del psicópata: algunas observaciones para entender la violencia depredadora humana”, en Raine & Sanmartín 2000: 43.

⁶³⁹ *Ibidem*: 44.

elaborando en su mente las asociaciones o redes semánticas y afectivas de las palabras que acababan de ver”.⁶⁴⁰ En general estos datos apuntan a la idea de que los psicópatas tienen serias dificultades para comprender el contenido emocional de una palabra, o al menos a la hora de reaccionar ante el mismo de forma apropiada.

En lo que respecta a estudios SPECT, los registros del flujo sanguíneo cerebral relativo (síntoma de una mayor actividad en el punto en que es registrado) durante un experimento⁶⁴¹ similar al descrito más arriba, revelaron que los psicópatas, a diferencia de los individuos normales (que usaban sobre todo la región anterior derecha del cerebro a la hora de analizar palabras con contenido emocional), activaban sobre todo la corteza occipital. Esto supone, dicho de forma algo vaga e imprecisa, que los psicópatas realizaban la tarea sin activar las áreas cerebrales asociadas al lenguaje. Para ellos la tarea no revelaba contenido semántico, se trataba tan sólo de una cuestión de reconocimientos léxicos que ponía en juego su atención, su capacidad visual, no su capacidad lingüística.

Por último, estudios con RMf⁶⁴² mostraron que “los no psicópatas exhibían mayor activación que los psicópatas durante el procesamiento de palabras con carga emocional en varias regiones límbicas, incluyendo la amígdala, íntimamente relacionada con la emoción, y el cíngulo, implicado en los procesos emocionales y de atención”.⁶⁴³ Estos datos explicarían la apatía característica de los psicópatas. Por otra parte, un sencillo experimento consistente, primero, en pulsar un botón a la vista de una X en un serie de

⁶⁴⁰ Ibidem.

⁶⁴¹ Intrator, J.; Hare, R.; Strizke, P.; Brichtswein, K.; Dorfman, D.; Harpur, T.; Bernstein, D.; Handelsman, L.; Schaefer, C.; Keilp, J.; Rosen, J. & Machac, J. 1997. “Brain imaging (SPECT) study of semantic and affective processing in Psychopaths”, *Biological Psychiatry* 42: 96-103.

⁶⁴² Kiehl, K. A.; Smith, A. M.; Mendrek, A.; Forster, B. B.; Hare, R. D. & Liddle, P. F. 1999. “Reduced limbic activity in criminal psychopaths during an affective memory task”, manuscrito en preparación.

⁶⁴³ Hare, R. D. “La naturaleza del psicópata: algunas observaciones para entender la violencia depredadora humana”, en Raine & Sanmartín 2000: 47.

preguntas, y en pulsarlo siempre que no apareciese una X, después, reveló que los psicópatas, aunque llevaban a cabo la tarea correctamente, apenas registraban actividad cortical a la hora de inhibir sus respuestas, y el caso es que “el control de la ejecución de las respuestas adecuadas y la inhibición de las respuestas inadecuadas reside en las regiones prefrontales ventromedial y dorsolateral”.⁶⁴⁴ De este experimento cabría deducir que en los psicópatas los mecanismos que facilitan el comportamiento adaptativo, las decisiones de tipo emocional y las reflexiones vinculadas a la toma de acciones, no funcionan correctamente. “Es decir, la conducta del psicópata no se ve modificada por las emociones que pueden ayudar a inhibir el comportamiento de una persona normal. Para ellos, los «frenos» emocionales del comportamiento (es decir, la conciencia) son débiles, y esto les permite cometer actos depredadores y violentos sin ningún escrúpulo.”⁶⁴⁵

En general se puede decir que todos estos experimentos vienen a indicar que “la discriminación automática entre estímulos con contenido emocional y neutral, que tiene lugar en los individuos normales (...) no se da en los psicópatas”.⁶⁴⁶ Pero los problemas emocionales de los psicópatas no terminan ahí. Otros experimentos muestran que los psicópatas reaccionan de forma más intensa que las personas normales ante escenas de víctimas. Este punto fue demostrado en un experimento⁶⁴⁷ en que se midió el sobresalto de psicópatas y no psicópatas a través de la intensidad del parpadeo provocado por un ruido inesperado. Esta intensidad es diferente en individuos normales ante imágenes desagradables (el sobresalto en presencia de un ruido inesperado es intenso), neutras (el sobresalto es menor) y agradables (el sobresalto está inhibido). Tanto las imágenes desagradables como las

⁶⁴⁴ Ibidem: 49.

⁶⁴⁵ Ibidem.

⁶⁴⁶ Patrick, C. J. “Emociones y psicopatía”, en Raine & Sanmartín 2000: 101.

⁶⁴⁷ Patrick, C. J.; Bradley, M. M. & Lang, P. J. 1993. “Emotion in the criminal psychopath: Startle reflex modulation”, *Journal of Abnormal Psychology* 102: 82-92.

agradables, y no las neutras, provocaban una inhibición del reflejo en los psicópatas: “los estímulos aversivos no provocaban una respuesta defensiva en los psicópatas. Estos resultados apoyan la hipótesis de que detrás de la psicopatía hay una falta de ansiedad o miedo (...)”.⁶⁴⁸ En otro experimento similar⁶⁴⁹ pero en que se distinguía, dentro de las imágenes desagradables, entre imágenes amenazadoras (por ejemplo un arma apuntando al espectador) e imágenes de víctimas (por ejemplo una agresión), se observó que se inhibía el sobresalto ante las escenas de víctimas y sólo se daba un leve sobresalto ante las escenas amenazadoras. Delinquentes no psicópatas, en cambio, sólo inhibían el sobresalto ante escenas agradables y registraban el mayor grado de sobresalto ante escenas amenazadoras. Así, la “inhibición del reflejo que tiene lugar en los psicópatas cuando están viendo escenas de víctimas indica que los psicópatas reaccionan ante estas escenas aversivas prestándoles gran atención”.⁶⁵⁰

En definitiva, como se dijo al principio de este apartado, los psicópatas son personas que, sin tener alterada la mente, esto es, siendo capaces de percibir correctamente la realidad y de funcionar de manera correcta desde el punto de vista cognitivo, no son en cambio capaces de comprender el significado emocional de muchas situaciones. Usando la metáfora de Wittgenstein: cuando se trata de emociones, si los psicópatas abren su caja para ver cómo es su escarabajo se encuentran con que está totalmente vacía.

La gestación social del psicópata

Bien, hasta aquí se han visto los hechos: hay gente que no forma parte de nuestra comunidad moral. Ahora, la pregunta relevante para el filósofo nos es tanto la de “¿cómo son estas personas?”, sino la de “¿cómo han llegado a

⁶⁴⁸ Patrick, C. J. “Emociones y psicopatía”, en Raine & Sanmartín 2000: 102.

⁶⁴⁹ Levenston, G. K.; Patrick, C. J.; Bradley, M. M. & Lang, P. J. En prensa. “The psychopath as observer: Emotion and attention in picture processing”, *Journal of Abnormal Psychology*.

ser lo que son?” En efecto, si resultase que la psicopatía es fruto exclusivamente de la genética, aquella no nos diría mucho sobre aquello en que consiste ser un sujeto libre y responsable, capaz de llevar a cabo una vida participante. En cambio, descubrir defectos de socialización en los psicópatas nos puede instruir acerca de las bases de la vida participante y desvelarnos sus componentes últimos. Ciertamente éstos ya han sido estudiados, se trata de: las emociones reactivas y las atribuciones de responsabilidad en general; las expectativas que les sirven de base; y el catálogo de excusas y exenciones que nos permiten suspender nuestros juicios morales. Así, lo que los factores psicosociales, presentes en la exclusión de algunos sujetos de una comunidad moral, podrían revelarnos, no es tanto cuáles son las bases del funcionamiento moral de dicha comunidad, sino más bien cuál es la genealogía de esas bases. Esto es de enorme interés porque permitiría naturalizar de forma no esencialista la vida participante, mostrándonos el darwiniano *cui bono* de la misma. Veamos pues cuáles son las raíces de la amoralidad de los que viven fuera de la comunidad moral.

Ya ha sido señalado repetidas veces que la *predisposición* biológica a la violencia es inútil sin la confluencia de factores ambientales que la tornen *disposición*, la cita que sigue es larga pero suficientemente significativa en este sentido:

[A]unque los hogares [de los asesinos en serie] tenían la apariencia exterior de normalidad, eran de hecho disfuncionales (...) La mitad [de los asesinos] tenían padres implicados en actividades criminales. Casi el setenta por ciento tenía una historia familiar de abuso de alcohol y consumo de drogas. Todos los asesinos –sin excepción– habían sufrido malos tratos emocionales en su infancia. Y todos evolucionaron hacia la condición que los psiquiatras califican como adultos sexualmente disfuncionales; es decir, incapaces de mantener una relación madura, entre iguales, con otro adulto (...). Las relaciones entre nuestros sujetos de estudio con sus respectivas madres fueron invariablemente frías, distantes, faltas de amor y cuidado (...). Crecieron en un entorno donde se les desatendía, y donde no se estableció límites para su comportamiento.⁶⁵¹

⁶⁵⁰ Patrick, C. J. “Emociones y psicopatía”, en Raine & Sanmartín 2000: 103.

⁶⁵¹ Ressler & Schachtman 1995: 100-101.

Robert K. Ressler, el autor de este texto, fue durante años investigador del FBI, resolvió numerosos casos de asesinos en serie y jugó un papel decisivo en la creación de la Unidad de Ciencias de la Conducta. El método seguido por él para trazar perfiles psicológicos de asesinos, la entrevista, ha sido adoptado también por psicólogos expertos en criminología para penetrar en los motivos de los asesinos en serie. Esas entrevistas en general han corroborado lo que se ha visto en los apartados anteriores: el asesino psicópata “por un lado, reconocía haber obrado mal –en el sentido de que sabía que había reglas sociales y legales que prohibían lo que había hecho-, pero, por otro, parecía estar convencido de que sus acciones estaban justificadas. Había, pues, una contradicción entre la conciencia cognitiva de la ilegalidad de su conducta, y el significado emocional de esa conducta.”⁶⁵² Además, las entrevistas mostraban que en general los asesinos en serie eran absolutamente maniqueos, manifiestamente incapaces de ver matices entre el tener razón y el no tenerla, y el que algo sea bueno o malo. Y aquí parece que se puede empezar a entender que, aunque dichas anomalías morales tengan la forma de lesiones cerebrales, tienen como sustrato una educación moral deficiente que no fue capaz de preparar a estos sujetos para participar adecuadamente en el mundo moral, para enfrascarse en “relaciones interpersonales normales”: “Parece que lo que nos distingue de ellos [los asesinos] no es nuestra naturaleza sino el significado que atribuimos a nuestras experiencias vitales.”⁶⁵³

La infancia es el momento de nuestras vidas en que los adultos nos enseñan, a veces de forma inconsciente y sutil, a convertirnos en seres participantes, en personas capaces de interpretar ciertas experiencias con carga moral en la misma línea en que lo harían la práctica totalidad de las demás personas. Así, gran parte de los fracasos a la hora de participar “en relaciones interpersonales normales”, se registran a causa de problemas en las relaciones

⁶⁵² Scrapeck C. A. “Los motivos del asesino en serie”, en Raine & Sanmartín 2000: 168.

⁶⁵³ Ibidem: 178.

del niño, a edades muy tempranas, con sus padres. A veces experiencias tan tempranas como las que tienen lugar después de las doce primeras horas de vida, hasta cuarenta y ocho horas más tarde, ya permiten que se perciba la influencia del ambiente en la personalidad, como muestra el siguiente experimento.⁶⁵⁴ A dieciocho bebés de sexo femenino sanos recién nacidos, se los separó en tres grupos y se les dejó en tres ambientes diferentes: en el primero de ellos “cuando el bebé estaba alerta, inquieto o llorando, normalmente entre toma y toma de leche, escuchaba una grabación de una madre que decía cosas como «hola, chiquitina, ¿cómo estás?»”⁶⁵⁵; en el segundo ambiente, cinco segundos después de la grabación un investigador cogía al bebé en brazos durante treinta y tres segundos y luego lo dejaba de nuevo en la cuna; en el tercer ambiente (grupo control) “a los bebés ni se les puso la grabación ni se les cogió en brazos durante las primeras sesenta horas de su vida”⁶⁵⁶. El resultado fue que los bebés del segundo grupo pasaban más tiempo despiertos y pasaban más tiempo llorando. “Los autores del estudio concluyeron: «Dado que estas niñas percibieron la relación entre estímulos auditivos (grabación) y el ser cogidas en brazos cuando lloraban, es probable que terminaran asociando el llanto con el ser cogidas en brazos» (...) En otras palabras, se les había enseñado a querer ser abrazadas y a llorar para comunicar este deseo.”⁶⁵⁷ Hasta nuestras respuestas aparentemente más instintivas ante estímulos ambientales están condicionadas socialmente, y en ausencia de ciertas condiciones sociales puede ocurrirnos que no aprendamos a responder adecuadamente a las situaciones que la vida nos obligue a enfrentar. “En los hogares normales los recién nacidos son cuidados por adultos que responden

⁶⁵⁴ Thoman, E. B.; Korner, A. F. y Benson-Williams, L. 1977. “Modification of responsiveness to maternal vocalization in the neonate”, *Child Development*, 48: 563-569. Es admirable cómo crece el conocimiento científico gracias al espíritu del juramento hipocrático.

⁶⁵⁵ McCord, J. “Contribuciones psicosociales a la violencia y la psicopatía”, en Raine & Sanmartín 2000: 211.

⁶⁵⁶ *Ibidem*: 212.

⁶⁵⁷ *Ibidem*.

al llanto o a los indicadores de hambre, dolor o cualquier otra necesidad. Este tipo de respuesta enseña a los bebés que su malestar puede ser aliviado por las acciones de otros.”⁶⁵⁸ En cierto sentido la vida en sociedad es una prueba de confianza, esa confianza en que los demás, en principio, tienen buenas intenciones para con uno. Dicha confianza se gesta precisamente cuando el niño aprende “que su malestar puede ser aliviado por las acciones de otros”, porque aprender esto es aprender que se puede confiar en los demás, en su buena voluntad. Cuando el niño no se mueve en esas situaciones de confianza se incrementa su estrés (por ejemplo, en las primeras revisiones médicas los bebés sufren bruscos incrementos de cortisol, señal de estrés). Experimentos con animales ⁶⁵⁹ han demostrado que el estrés en edades tempranas puede acabar conduciendo a la muerte celular en partes del cerebro, y en cualquier caso afecta al desarrollo de áreas cerebrales relacionadas con la memoria, las emociones negativas y la regulación de la atención. En fin, muchas de aquellas que, como se vio más arriba, presentan serios daños en los psicópatas. La forma de proteger, según los investigadores implicados en el citado experimento, el desarrollo cerebral del niño y la reactividad posterior al estrés es la presencia de “«[c]uidadores sensibles y responsables, y relaciones afectivas seguras en la infancia y la niñez»”.⁶⁶⁰ En fin, un entorno en que sea posible la confianza, en que tenga sentido tener la expectativa de que los demás, siempre que les sea posible, evitarán nuestro sufrimiento. Este “siempre que les sea posible” no está exento de importancia, porque una de las circunstancias en que no es posible evitar un cierto daño (por supuesto no forzosamente físico) al niño, es cuando el tutor considera que éste se ha comportado mal y entonces lo castiga. El castigo en la educación moral es indispensable, y los

⁶⁵⁸ Ibidem.

⁶⁵⁹ Gunnar, M. R. y Barr, R. G. 1998. “Stress, early brain development, and behavior”, *Infants and Young Children*, 11 (1): 1-14.

⁶⁶⁰ McCord, J. “Contribuciones psicosociales a la violencia y la psicopatía”, en Raine & Sanmartín 2000: 213.

males que una forma errónea e injusta de usar esa herramienta puede llegar a producir son proporcionales a la importancia de la misma. Los castigos arbitrarios, desproporcionados y continuos no permiten al niño aprender a tener expectativas adecuadas, esto es, no le permiten aprender a esperar que cuando haga algo bien será premiado y que cuando haga algo mal será castigado.

Por otra parte el castigo físico puede llegar a convertirse en algo inútil si los niños llegan a desarrollar, como les pasa a los psicópatas, una alta tolerancia al dolor. Algunos niños, aquellos que demuestran haber recibido poca atención a lo largo de su vida, presentan, en el sentido de una mayor capacidad de aguante, una escasa sensibilidad al dolor (un caso concreto reflejado en un estudio acerca de este tema⁶⁶¹ habla de una niña que tuvo durante dos días una astilla metálica en la córnea y no avisó a ningún adulto para que le fuera retirada). Estos niños no tienen ningún problema en las terminaciones nerviosas que no les permita sentir dolor, sencillamente “en un entorno insensible, estos niños habían aprendido a ignorar su propio sufrimiento y probablemente también el dolor ajeno. Esta falta de empatía es una de las características destacables en los psicópatas.”⁶⁶² Existe un código universal para solicitar la atención de los demás: el gesto facial.⁶⁶³ Los niños aprenden a mostrar tristeza, miedo y a sonreír para lograr la atención de los adultos:

Sin embargo, cuando los estímulos no son los adecuados, como ocurre en las familias con padres abusivos y negligentes, los niños sonríen poco, con lo cual aún reciben menor atención. Esto conlleva a que difícilmente aprendan aquellas conductas que tienden a producir

⁶⁶¹ Goldfarb, W. 1945. “Psychological privation in infancy and subsequent adjustment”, *American Journal of Orthopsychiatry*, 15: 247-255, y Goldfarb, W. 1958. “Pain reactions in a group of institutionalized schizophrenic children”, *American Journal of Orthopsychiatry*, 28: 777-785.

⁶⁶² McCord, J. “Contribuciones psicosociales a la violencia y la psicopatía”, en Raine & Sanmartín 2000: 215.

⁶⁶³ Véanse a este respecto Ekman, P. y Oster, H. 1979. Traducción “Expresiones faciales de la emoción”, en *Estudios de Psicología* (1981): 117-139, e Iglesias, J.; Naranjo, J. M^a; Picazo, C. y Ortega, J. E. 1984. “La cara y la emoción: datos para una réplica”, *Estudios de Psicología*, 18: 101-109.

reacciones positivas en los demás. Hasta cierto punto, los niños aprenden a saber cómo se sienten los demás a través de las expresiones físicas y verbales, pero los padres que abusan y son negligentes dedican poco tiempo a comunicarse con sus hijos (...). De ahí que lo más probable sea que estos niños apenas comprendan qué piensan y qué sienten los demás.⁶⁶⁴

Así, la falta de atención puede convertir a un niño en un *autista moral*. Y no sólo eso. Además de al estrés y a la falta de reacción ante el dolor propio y ajeno, el abandono y la dejadez de los padres a la hora de atender a sus hijos puede llevar a éstos a altos índices de agresividad. En efecto, si los castigos son arbitrarios y el niño ve que salvo por medio de la violencia es imposible llamar la atención de sus progenitores, acabará concibiendo la agresividad como algo positivo. “[L]os niños tendrían que estar dispuestos a renunciar a los beneficios que comporta su conducta agresiva, a cambio de los beneficios que se obtienen al formar parte de una unidad social”⁶⁶⁵, pero cuando esa unidad social está rota y a los niños no se les permite obtener ningún beneficio de sus “relaciones interpersonales”, entonces éstos aprenden a considerar a los demás como medios con los cuales es lícito utilizar la violencia, ya que de ninguna otra forma podrían obtener algo favorable de ellos.

Relacionando lo que más arriba dije acerca del castigo y esto último que se acaba de ver acerca de la agresividad, se obtiene otro dato importante: la exposición temprana al castigo *severo* predispone al comportamiento criminal violento. Un estudio⁶⁶⁶ hecho a lo largo de cinco años sobre un total de doscientos chicos, con el objeto de ver los efectos del castigo físico cuando éste se descubría injustificado, esto es, no provocado por el mal comportamiento del niño, confirmó la existencia de esa predisposición a la violencia como resultado del castigo severo. En dicho estudio se visitaron cada dos semanas los hogares de todos esos chicos y se midieron el afecto y cariño paternos, y la intensidad de los castigos que se les aplicaba. Treinta años más

⁶⁶⁴ McCord, J. “Contribuciones psicosociales a la violencia y la psicopatía”, en Raine & Sanmartín 2000: 215.

⁶⁶⁵ Ibidem: 216.

⁶⁶⁶ McCord, J. 1997. “On discipline”, *Psychological Inquiry*, 8 (3): 215-217.

tarde se localizó a los chicos que protagonizaron este estudio y se comprobó sus antecedentes penales. “Los análisis indicaron que el castigo corporal infligido por sus madres había aumentado la probabilidad de que fueran condenados por delitos violentos, incluso cuando hubo afecto materno.”⁶⁶⁷ Además, otros estudios han demostrado que el castigo severo en lugar de hacer que el niño pierda interés por lo prohibido provoca el efecto contrario. En uno de estos estudios⁶⁶⁸ un grupo de niños tenía que ordenar cinco juguetes según sus preferencias, y a cada uno de estos niños se le prohibió jugar con el segundo de su lista personal. Se les separó en dos subgrupos: en uno de ellos jugar con el juguete prohibido conllevaba una “amenaza ligera”, a saber, el investigador amenazaba con enfadarse mucho si el niño jugaba con el juguete prohibido; en el otro subgrupo, jugar con el juguete prohibido iba acompañado de una “amenaza severa”, a saber, el investigador amenazaba con enfadarse mucho y llevarse todos los juguetes si el niño jugaba con el prohibido. A los cuarenta y cinco días, cuando a los niños se les pidió que reordenaran los juguetes según sus preferencias, no sólo ninguno de los niños de la condición “amenaza severa” había perdido el interés por el juguete prohibido (cuando algunos de la condición “amenaza ligera” sí lo habían hecho), sino que una gran parte de ellos lo colocaba ahora en el primer lugar de su ranking de juguetes.

Por último hay que decir que no es indispensable ser objeto de violencia para malograrse moralmente, a menudo basta con vivir expuesto a ella: “los conflictos conyugales están fuertemente relacionados con el posterior comportamiento violento de los hijos”⁶⁶⁹.

⁶⁶⁷ McCord, J. “Contribuciones psicosociales a la violencia y la psicopatía”, en Raine & Sanmartín 2000: 217.

⁶⁶⁸ Aronson, E. y Carlsmith, J. M. 1963. “Effect of the severity of threat on the devaluation of forbidden behavior”, *Journal of Abnormal and Social Psychology*, 66 (6): 584-588.

⁶⁶⁹ McCord, J. “Contribuciones psicosociales a la violencia y la psicopatía”, en Raine & Sanmartín 2000: 218.

En definitiva, puede decirse que para dotarnos con un sistema de valores deben darse ciertas condiciones sociales que nos permitan interiorizar las reglas básicas de la vida participante, de la comunidad moral. La falta de atención por parte de los padres, la violencia en el entorno familiar, la arbitrariedad de los castigos y su desproporción, favorecen que se dé la ausencia de empatía que caracteriza al psicópata y que lo sitúa fuera de la comunidad moral. Resta preguntarse por todo lo contrario: ¿Qué es lo que nos hace pertenecer a la comunidad moral, cómo nos dotamos de un sistema de valores? Algunas partes de la respuesta a esta pregunta han ido apareciendo en páginas precedentes, pero falta una teoría integral que permita entender cómo nos hacemos responsables areteicamente. Mi hipótesis es que nos hacemos responsables areteicamente del mismo modo en que aprendemos a “hacer cosas con palabras”.

III.4.3. Aprender a participar

“La Teoría del Constructo (*Construct Theory*) establece que la motivación surge cuando enseñamos a los niños a justificar sus acciones mediante razones.”⁶⁷⁰ Esta teoría ha sido desarrollada por la profesora de criminología Joan McCord (entre cuyos méritos está el ser la vicepresidenta de la Sociedad Internacional de Criminología de Norteamérica), con el objeto de explicar por qué los psicópatas y otros delincuentes cometen actos violentos. Así pues, dicha teoría defiende algo similar a lo que se viene proponiendo aquí, a saber: que, para que una persona se convierta en una persona normalmente responsable (cosa esta que no son los psicópatas), es necesario que se la enseñe a participar del juego de dar razones para las acciones que le permitan establecer juicios sobre expectativas, esto es, juicios de carácter moral sobre

acciones responsables. “(...) [L]os motivos relacionan las razones facilitadoras con la acción. Estas razones facilitadoras son a las acciones lo que los argumentos a las creencias.”⁶⁷¹ ¿Y de dónde surgen esas razones? La hipótesis de McCord concuerda una vez más con el espíritu de este trabajo: “(...) el lenguaje tiene fundamentalmente un carácter evaluativo y público. Esta posición evita los problemas relacionados con las teorías de la acción que insisten en que hay un abismo entre las descripciones y las evaluaciones. También se evitan los problemas derivados de las teorías del lenguaje que colocan los significados en la esfera privada de las imágenes mentales.”⁶⁷²

La idea de McCord es que aprender a usar el lenguaje es aprender qué se debe creer y qué se debe hacer, pues el lenguaje nos permite clasificar el mundo, y por tanto “que los motivos forman parte de una clase especial de descripciones y que son razones que facilitan la acción”.⁶⁷³ Así, al aprender el lenguaje un niño no sólo está aprendiendo qué es cada cosa, sino qué es digno de atención y qué no lo es, qué es deseable y qué indeseable, qué es grato y qué es peligroso: “Al aprender como nombrar y reidentificar objetos y actividades, los niños construyen clasificaciones. Los sistemas de clasificación que desarrollen determinarán su modo de percibir el mundo y desenvolverse en él.”⁶⁷⁴ Así, los educadores, al mismo tiempo que enseñan un lenguaje, están enseñando estas clasificaciones, están definiendo el mundo moral del niño: “la socialización influye en el comportamiento porque determina las razones potenciadoras o facilitadoras que constituyen el motivo esencial para llevar a cabo una acción determinada.”⁶⁷⁵ Por ello la responsabilidad personal no depende de nadie, pero la responsabilidad areteica es algo que se aprende a tener. Depende de nuestros educadores que lleguemos a poseer una teoría de

⁶⁷⁰ Ibidem: 219.

⁶⁷¹ Ibidem: 220.

⁶⁷² Ibidem.

⁶⁷³ Ibidem.

⁶⁷⁴ Ibidem: 221.

la virtud base, una “habilidad para actuar de acuerdo con lo Verdadero y lo Bueno”⁶⁷⁶, un sistema de valores:

[E]l hecho de que diferentes personas valoremos un mismo suceso como deseable o indeseable dependiendo de nuestro *sistema de valores* [se explica por las razones facilitadoras que nos fueron transmitidas de niños]. Es en la infancia cuando los cuidadores proporcionan los acontecimientos sobre los que el niño crea su sistema de categorías para clasificar su entorno social. Al enseñar a sus hijos cómo deben actuar, los padres transmiten una serie de señales que indican su propio sistema de valores. Algunos padres enseñan a sus hijos las consecuencias del mal comportamiento en términos del daño que provocan a los demás o a los objetos. Estas técnicas «inductivas» parecen ser muy eficaces a la hora de reducir las probabilidades de que un niño se comporte mal en el futuro (...).⁶⁷⁷

Una vez más: el castigo. El castigo no transmite tan sólo patrones de comportamiento desnudos de significado, sino que enseña a significar, enseña a “puntuar” acciones, lo cual viene a ser tanto como aportar razones a favor o en contra de las mismas. En fin, el castigo suministra un sistema de valores porque nos enseña a entender el lenguaje moral. “El aprendizaje del lenguaje requiere asimilar lo que J. L. Austin llamó los «*performatives*» [realizativos], o expresiones de petición, disculpa, promesa, aviso y recomendación. Los niños aprenden estas expresiones incluso antes de aprender a describir lo que ven y lo que oyen.”⁶⁷⁸ Con razón dice Austin que “examinar las excusas es examinar casos en que ha habido alguna anormalidad o fallo: y como sucede tan a menudo, lo anormal podría arrojar luz sobre lo normal, podría ayudarnos a penetrar el velo de facilidad y obviedad que esconde los mecanismos del acto naturalmente exitoso.”⁶⁷⁹

Según Austin, cuando un agente hace algo malo, hay básicamente dos formas en que puede defender su conducta. La primera es justificando la acción: “admitiendo francamente que él, X, hizo realmente esa cosa, A, pero argumentar que ello era bueno, o correcto, o sensato, o algo que está

⁶⁷⁵ Ibidem: 224.

⁶⁷⁶ Wolf 1990: 79.

⁶⁷⁷ McCord, J. “Contribuciones psicosociales a la violencia y la psicopatía”, en Raine & Sanmartín 2000: 222. La cursiva es mía.

⁶⁷⁸ Ibidem: 221.

⁶⁷⁹ Austin 1956-7, en Urmson & Warnock 1961: 180.

permitido hacer, tanto en general como en las circunstancias especiales de esa ocasión.”⁶⁸⁰ La segunda explicando que su acción no fue tan suya como podría parecer en un principio: “admitiendo que aquello que se hizo no era algo bueno, pero argumentar que no es justo o correcto decir *escuetamente* ‘X hizo A’.”⁶⁸¹ Y añade Austin: “En una defensa, resumiendo, aceptamos la responsabilidad pero negamos que fuese algo malo; en la otra, admitimos que fue malo pero no aceptamos plena, o incluso ninguna, responsabilidad.”⁶⁸²

El que alguien sea capaz de distinguir entre justificarse y excusarse supone que es capaz de distinguir entre *hacer* algo malo y *ser culpado* por algo malo. Puede decirse que precisamente los psicópatas no son capaces de hacer esta distinción. Así, cuando un psicópata trata de excusarse, lo que cree estar haciendo es justificarse, porque para él el límite del bien y del mal es que le culpen o que no le culpen, que le castiguen o que no lo castiguen. En efecto, si una persona es sometida durante la infancia a castigos arbitrarios continuos, ésta será incapaz de llegar a asociar el castigo a un determinado patrón de comportamiento que correspondería a aquello que llamamos “hacer algo malo”. Veamos un ejemplo. Imaginemos que un niño no deseado, cuyos padres son alcohólicos e irascibles, recibe una paliza si estudia y saca buenas notas porque eso supone, según sus progenitores, que está holgazaneando con los libros y no ayuda en las tareas de la casa, o que desprecia a sus padres que no tienen estudios, o que no colabora suficientemente en el negocio familiar. E imaginemos que también recibe una paliza si no estudia tanto y saca peores notas porque eso supone, una vez más según sus progenitores, que es un holgazán que se pasa los días sin abrir un libro, un ingrato que no valora lo que tiene, o un imbécil que es incapaz de ver la importancia de la escuela y no piensa más que en ganar dinero rápido en vez de mejorar en la escala social.

⁶⁸⁰ Ibidem: 176.

⁶⁸¹ Ibidem.

⁶⁸² Ibidem.

En fin, las razones que se le pueden ocurrir a esos maltratadores para castigar a su hijo son innumerables, porque no requieren ser coherentes entre sí, son absolutamente arbitrarias. En cualquier caso, poniéndonos en la perspectiva del niño, si “no estudiar” es algo por lo que me pueden castigar, y “estudiar” es también algo por lo que me pueden castigar, sin que sea posible determinar en qué condiciones es correcto hacer lo uno y en qué condiciones hacer lo otro, nunca será posible que ese niño llegue a atribuirle valor en sí misma a la conducta “estudiar”, sólo lo hará *a posteriori*, en función de ser o no castigado. Así, su vida será sin expectativas, pues tener expectativas supone juzgar *antes de*, pero el psicópata sólo puede juzgar en función de que le cojan y le condenen, o no, o sea, siempre *después de*. O tal vez sea correcto decir que el psicópata ni tan siquiera puede juzgar moralmente, pues si no se le ha enseñado a distinguir entre justificación y excusa, no se le ha enseñado a distinguir entre ser responsable y no serlo. Por ello el psicópata podrá usar sintácticamente todos los términos vinculados a la responsabilidad, desde las emociones morales hasta las exenciones, pero no les adjudicará ningún contenido semántico. La responsabilidad, como se vio al final de la segunda parte de este trabajo, es una forma de vida. En esa forma de vida se entra participando con los demás de las reglas de ese juego lingüístico que es el mundo participante: “Del mismo modo que la capacidad de un hablante competente para usar un lenguaje para expresar sus propios pensamientos, un agente moralmente competente es capaz de formar sus propios motivos moralmente impondibles sólo si posee un dominio suficiente de las convenciones que asignan interpretaciones de motivos-tipo a acciones-tipo.”⁶⁸³ El psicópata es alguien a quien no se le ha enseñado a participar en “relaciones interpersonales normales”, de tal forma que no domina esas convenciones para asignar interpretaciones. Pero lograr interpretar

⁶⁸³ McKenna 1998b: 6-7. Relaciónese esta cita con lo dicho por McCord en las citas 671 y 673.

correctamente, aprender a participar, es lo mismo que convertirse en una persona responsable, en un ser libre. Recuérdese que para que sea posible hacer juicios de responsabilidad, tanto emotivos como reflexivos, es necesario poseer cierta actitud proposicional consistente en la evaluación de una demanda (que si es moral es una obligación). Así, para entender qué es ser responsable hay que comprender primero que existen demandas, algunas de ellas obligaciones, que es necesario satisfacer y que nos sean satisfechas. De la educación depende que a alguien se le dote con un sistema de valores, lo cual no es sino un ranking de demandas. De la educación depende pues que alguien comprenda en qué consiste ser responsable, lo cual equivale a serlo. La libertad es algo que se aprende.

Recapitulación

Sin duda he dado muchas vueltas y debo hacer una pausa (final, por cierto) para engarzar todo aquello que se ha ido viendo hasta ahora. Hemos aprendido algunas cosas acerca del concepto de responsabilidad moral analizando el artículo de Strawson “Freedom and Resentment”. De su análisis y posterior comentario se pueden sacar aquí las siguientes conclusiones.

El ser humano vive inmerso en relaciones interpersonales que llevan consigo una actitud participante consistente precisamente en nuestra actitud normal a la hora de relacionarlos con los demás. Esa vida participante incluye dos tipos de actitud a la hora de evaluar las acciones propias y ajenas: la reactiva y la reflexiva. La primera se compone principalmente de atribuciones emotivas (mediante emociones reactivas), aunque también de disculpas; la segunda de exculpaciones reflexivas (exenciones y excusas), pero también veredictos de culpabilidad. Las emociones reactivas son emociones que se dan en nosotros cuando creemos que ciertas expectativas que tenemos puestas en los demás y en nosotros mismos se han visto o frustradas o cumplidas. En estas emociones se basa nuestro concepto de responsabilidad, pues dichas emociones se articulan mediante dicho concepto. Dentro de la actitud reactiva a veces interactuamos con los demás moralmente y a veces no, todas las emociones reactivas pueden estar presentes en ambas situaciones salvo la indignación y la culpa, que son específicamente morales, de tal forma que en ellas se basa nuestra responsabilidad moral. Frente a la actitud reactiva podemos, en ciertas ocasiones, adoptar una actitud reflexiva. Esta actitud reflexiva consiste en lo siguiente: por un lado en una investigación teleológica, y por otro en un intento por obtener cierto resultado de aquel a quien dispensamos esta actitud. Esta actitud sólo se ve justificada en ciertos casos (en que hay exenciones o excusas para la responsabilidad de aquel a quien

juzgamos), pero a veces puede ser adoptada en otros (en aquel contexto que requiere que las emociones estén presentes). Ocupándome de los casos en que se ve justificada he trazado una tabla de excusas y exenciones que viene a confirmar la hipótesis 5) de la presentación de este trabajo, a saber: “Tenemos un mapa de excepciones que nos permite suspender esos juicios de responsabilidad.”

Me he ocupado también, para ahondar en la comprensión de la actitud reflexiva y del sistema de excusas y exenciones, del alcance de la cultura en nuestro sistema emotivo de atribución de responsabilidad, y he terminado haciendo hincapié en la absoluta necesidad de que los agentes cuenten con y sean capaces de actuar de acuerdo a un sistema de valores culturalmente constituido. Así, se puede sacar la conclusión de que la responsabilidad moral se *expresa* en nuestras emociones reactivas, pero se basa en algo más: en un sistema de valores compartido que configura una comunidad moral. Ese sistema de valores es lo que nos permite tener ciertas expectativas en que se basan nuestras emociones reactivas, y es por lo tanto el que nos permite hablar de responsabilidad moral.

Por último he estudiado cuáles son las condiciones necesarias para que sea adecuado tener expectativas sobre alguien y por ende considerarle un agente moralmente responsable: que posea *responsabilidad personal* y *responsabilidad areteica*. Afirmar esto es lo mismo que decir lo siguiente: toda persona, para ser juzgada moralmente responsable, ha de actuar voluntariamente (a) con conocimiento de causa, (b) sin coacción y (c) siendo miembro de una comunidad moral. Quien no cumpla estos requisitos debe ser excusado o eximido de responsabilidad moral.⁶⁸⁴ Así, parece que aquello en

⁶⁸⁴ No forzosamente debe ser eximido o excusado totalmente, a veces las excusas y exenciones no constituyen más que atenuantes. Sería interesante estudiar estos matices, pero por desgracia habrá que conformarse aquí con los trazos que he sido capaz de dar.

que consistiría a grandes rasgos la responsabilidad moral sería lo siguiente: la capacidad para actuar libremente en una comunidad moral.

Y bien, una vez hecha esta recapitulación, creo que ya puedo permitirme terminar este trabajo resumiendo sus conclusiones más importantes.

CONCLUSIÓN

Llega por fin el momento de resumir la teoría filosófica de la libertad que he tratado de construir en las páginas anteriores. Un buen método para hacerlo puede ser recordar las cinco hipótesis que he ido corroborando a lo largo del trabajo, y ver cómo encajan unas con otras:

- 1) Las acciones humanas están determinadas, determinadas por los propios seres humanos.
- 2) Una o más personas pueden actuar de forma diferente en situaciones muy parecidas.
- 3) Es posible predecir comportamientos, sin embargo las acciones humanas poseen un buen grado de imprevisibilidad.
- 4) Tenemos expectativas y emitimos juicios de responsabilidad sobre las acciones ajenas y propias basándonos en dichas expectativas.
- 5) Tenemos un mapa de excepciones que nos permite suspender esos juicios de responsabilidad.

En efecto, en la primera parte del trabajo se vio que para tener libertad de la voluntad, el ser humano ha de contar con control sobre sus acciones y con la posibilidad de actuar de otro modo. La primera hipótesis, que ya cobra aquí, como todas las demás, estatuto de conclusión, confirma el hecho de que los seres humanos controlan sus acciones pues, dado que existe el determinismo y que los agentes son entidades deterministas autoconscientes, entonces son capaces de determinar sus acciones, lo cual viene a ser controlarlas.

Se vio también, en esa primera parte, que la idea de que los agentes operan controlando determinísticamente sus acciones, suele chocar, aunque realmente no es así, con la idea de que los seres humanos podrían actuar de otro modo que como lo hacen. Sin embargo, dicha capacidad de contar con posibilidades alternativas no es incompatible con el determinismo, pues no tiene sentido si es comprendida como que una misma persona, en exactamente una misma situación, sería capaz de llevar a cabo dos acciones distintas. Pero sólo en ese sentido las posibilidades alternativas contradicen el determinismo. Si por el contrario, se entiende que contar con posibilidades alternativas es, como reza la segunda hipótesis/conclusión, “actuar de forma diferente en situaciones muy parecidas”, entonces no sólo el determinismo no supone una amenaza para las posibilidades alternativas, sino que puede ser su fundamento.

En lo que respecta a la tercera hipótesis/conclusión, ésta hace referencia al miedo que tienen los detractores del determinismo a la predecibilidad que conllevan los sucesos naturales determinados. Lo que he tratado de mostrar es que, efectivamente, los comportamientos humanos no por ser libres son impredecibles, pero que si se le pierde el miedo a un Gran Predictor Universal de comportamientos, entonces uno se da cuenta de que las acciones humanas, cuando son analizadas por otros seres humanos, suelen ser suficientemente impredecibles como para dar cuenta de esa idea más o menos intuitiva de que, si somos libres, entonces no se puede hacer con nosotros algo parecido a una “ley de la caída de los graves”.

En fin, la primera parte del trabajo lleva a la conclusión de que existe la libertad de la voluntad, y que ésta incluye determinación consciente de sus acciones por parte del agente, posibilidad de actuar de forma diferente en situaciones similares y cierta impredecibilidad dada la riqueza de los motivos que mueven al agente a actuar. Y he aquí una palabra clave: motivos.

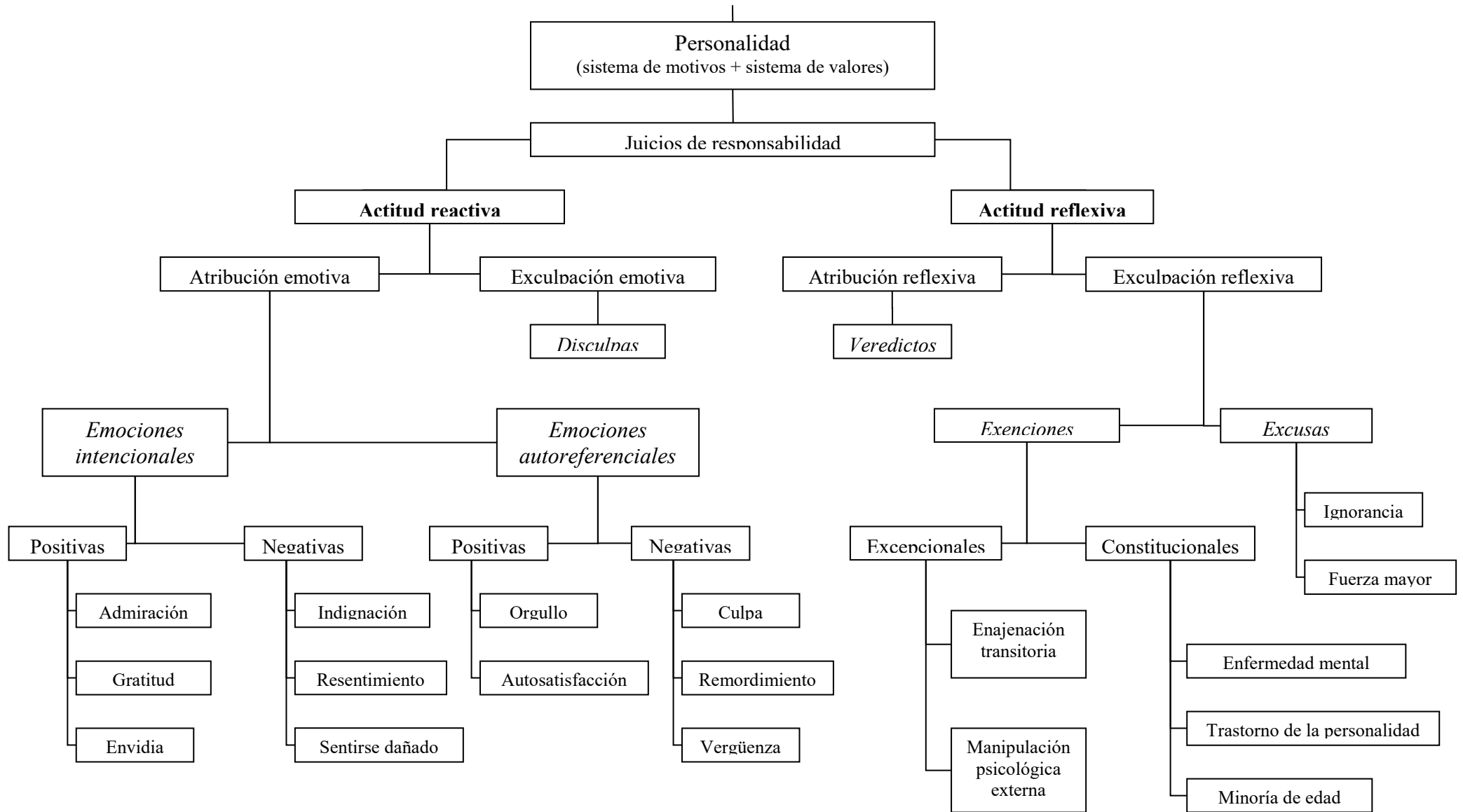
En la segunda y tercera partes del trabajo me he ocupado de ver cómo se articulan socialmente esos motivos que componen el carácter del agente, la

base de su personalidad. He descubierto entonces que la personalidad del agente, con sus motivos, es fruto de la libertad de la voluntad aunque no basta para la responsabilidad moral. Pero la responsabilidad es precisamente la consecuencia más importante de la posesión de la libertad. El complemento del carácter dentro de la personalidad del agente, que nos convierte en seres responsables, son ciertas expectativas que nos permiten evaluar nuestras acciones y las de los demás, unas expectativas que componen un sistema de valores que, si se adjunta al sistema de motivos, al carácter, da como resultado un agente responsable. Por ello, confirmar la cuarta hipótesis (según la cual “tenemos expectativas y emitimos juicios de responsabilidad sobre la acciones ajenas y propias basándonos en dichas expectativas”) como hice en la segunda parte del trabajo, viene a ser confirmar que los seres humanos somos seres responsables.

En lo que respecta a la última hipótesis, el concepto mismo de responsabilidad conlleva que pueda considerarse que alguien es responsable de sus acciones precisamente porque podría no serlo, esto es: la responsabilidad no incluye tan sólo atribuciones, sino también exculpaciones. Por ello poseemos un “mapa de excepciones” según el cual somos capaces de decidir cuándo alguien ha actuado responsablemente y cuándo no, porque sin dicho mapa no podría existir la forma de vida en que consideramos a las personas libres y responsables.

Así, la confirmación de estas cinco hipótesis viene a ser el fundamento de una teoría de la libertad de la voluntad y el agente responsable, pero no nos dice mucho acerca de sus pormenores. Por ello, sobre todo en la tercera parte, me ocupé de observar nuestra vivencia de la responsabilidad moral intentado entresacar de ella sus rasgos característicos. En la página siguiente presento un esquema que pretende aunar dichos rasgos y que resume la teoría desarrollada en este trabajo:

MUNDO PARTICIPANTE



A grandes rasgos, lo que se ha podido ver en la última parte del trabajo es que aquello en que consiste ser libre y responsable consiste sencillamente en lo siguiente: pertenecer a una comunidad moral, en ser consciente de las consecuencias de las propias acciones y en no estar coaccionado a la hora de actuar.

Pertenecer a una comunidad moral es contar con un sistema de valores que actúe conjuntamente con el sistema de motivos que nos impulsa a actuar. Contar con ese sistema de valores consiste en dominar un juego de expectativas según el cual somos capaces de evaluar nuestras acciones y las de los demás en razón de si cumplen o no ciertas expectativas que todos legítimamente podemos tener a la hora de relacionarnos unos con otros. En general todos esperamos de los demás cierta “buena voluntad”, esto es, podemos cabalmente esperar de cualquiera que pertenezca a nuestra comunidad moral que no nos dañará siempre que no cuente con alguna justificación que sea adecuada, esto es, que entre dentro del código de excusas de nuestra comunidad moral. El código de excusas de nuestra comunidad moral se rige por los otros dos requisitos que he presentado como fundamentales para la responsabilidad, a saber: no actuar ignorando lo que se está haciendo y no actuar coaccionado.

Cuando un miembro de nuestra comunidad moral cumple o no una de nuestras expectativas provoca en nosotros un cierto juicio sobre responsabilidad que las más de las veces tiene la forma de una emoción reactiva. Según el tipo de expectativa, según quién sea el agente sobre el que se vuelca esa expectativa, y según si nuestra expectativa se haya visto o no cumplida, sentiremos diferentes emociones reactivas. Básicamente sentiremos indignación o resentimiento si alguien viola una de nuestras expectativas, y admiración o gratitud si alguien la cumple. Si somos nosotros mismos quienes defraudamos nuestra propia expectativa sentiremos culpa o remordimiento, y orgullo o autosatisfacción si por el contrario respondemos a la demanda que

nosotros mismos nos hayamos hecho. También es posible hacer atribuciones de responsabilidad no emotivamente, así sucede con los veredictos que tienen lugar en los juicios.

Por otra parte, si un miembro de nuestra comunidad moral actuó ignorando lo que hacía, o forzado por algo o alguien ajeno a su propia voluntad, entonces consideraremos que cuenta con una buena excusa para su violación de nuestra expectativa y suspenderemos nuestras emociones reactivas. Y si llegamos a saber de alguien que ha violado una de nuestras expectativas que está enajenado, o que lo estaba cuando lo hizo, sabremos que no pertenece a nuestra comunidad moral, que está eximido de responsabilidad moral, y que entonces tampoco es un objeto adecuado para nuestras emociones reactivas.

Este juego de atribuciones de responsabilidad, de excusas y exenciones conforma nuestra vida social, nuestro mundo. Se trata de un mundo participante en el que hay que seguir ciertas normas que nos dicen qué vale y qué no vale para atribuir o no responsabilidad y para pedir o no cuentas por determinadas acciones a los demás o a uno mismo. El trasfondo de ese mundo es la libertad, una libertad entendida como responsabilidad tal y como acabo de definirla en los párrafos anteriores, y no contar con ese trasfondo no nos permite vivir correctamente en el mundo participante, no nos permite jugar correctamente ese juego.

Hay dos formas de jugar mal ese juego, de estar fuera de él: no perteneciendo a una comunidad moral, y siendo un incompatibilista y ejerciendo en el día a día como tal. Esto no quita que pueda llegar algún día en que cambiemos de mundo y de juego, pero mientras ese día no llegue *lo más racional* es comportarnos como habitantes del mundo en que efectivamente vivimos, esto es, comportarnos como los seres moralmente responsables que somos.

Ser libre es vivir en un mundo participante. Así, hoy por hoy, ser libre es algo inevitable.

BIBLIOGRAFÍA

- Agustín de Hipona. 1990. *Confesiones*, Madrid: Alianza.
- . 1971. *Obras de San Agustín III*, Madrid: BAC.
- Algra, K. Barnes, J. Mansfeld, J. Schofield, M. eds. 1999. *The Cambridge History of Hellenistic Philosophy*, Cambridge: Cambridge University Press.
- Anscombe, G. E. M. 1956. "Aristotle and the Sea Battle", *Mind* Vol. 65, No. 257: 1-15.
- Anselmo de Aosta. 1952. *Obras de San Anselmo I*, Madrid: BAC.
- . 1953. *Obras de San Anselmo II*, Madrid: BAC.
- Aristóteles. 1932. *Obras completas VII: Gran Ética, República ateniense, Economía*, Madrid: Espasa-Calpe.
- . 1998. *Ética nicomáquea. Ética eudemia*, Madrid: Gredos.
- . 1999. *Categorías. De la interpretación*, Madrid: Tecnos.
- . 2000. *Acerca del alma*, Madrid: Gredos.
- Arteta, Aurelio. 2002. *La virtud en la mirada. Ensayo sobre la admiración moral*. Valencia: Pre-Textos.
- Austin, J. L. 1956. "Ifs and Cans", *Proceedings of the Aristotelian Society*, reimpresso en Urmson & Warnock, 1961: 153-180.
- . 1956-7. "A Plea for Excuses", *Proceedings of the Aristotelian Society*, reimpresso en Urmson & Warnock, 1961: 175-204.
- Ayer, Alfred J. 1954. "Freedom and Necessity", en Ayer, *Philosophical Essays*, London: Macmillan. Reimpresso en Watson, 1982: 15-23.
- . 1980. "Free Will and Rationality", en Van Stratten, 1980: 1-13.
- Baudelaire, Charles. 1991. *Las flores del mal* (edición bilingüe), Madrid: Cátedra.
- Bennett, Jonathan. 1980. "Accountability", en Van Stratten, 1980: 14-47.
- Benson, Paul. 1990. "The Moral Importance of Free Action", en *Southern Journal of Philosophy* 28: 1-18.
- Bergson, Henry. 1999. *Ensayo sobre los datos inmediatos de la conciencia*, Salamanca: Sígueme.
- Berofsky, Bernard. ed. 1966. *Free Will and Determinism*, New York: Harper and Row.
- . 1987. *Freedom from Necessity: The Metaphysical Basis of Responsibility*, New York: Routledge & Kegan Paul.
- Bertomeu, M. J., Gaeta, R. & Vidiella, G. comps. 2000. *Universalismo y Multiculturalismo*, Buenos Aires: Eudeba.
- Boecio, Severino. 1999. *La consolación de la filosofía*, Madrid: Alianza.
- Borges, Jorge Luis. 1971. *Ficciones*, Madrid: Alianza.
- Brandom, Robert B. 2000. *Articulating Reasons. An introduction to Inferentialism*, Cambridge (Mass.): Harvard University Press.

- . 2002. *Tales of the Mighty Dead: Historical Essays in the Metaphysics on Intentionality*, Cambridge (Mass.): Harvard University Press.
- Bratman, Michael E. 1999. *Faces of Intention: Selected Essays on Intention and Agency*, Cambridge: Cambridge University Press.
- Broad, C. D. 1952. "Determinism, Indeterminism and Libertarianism", reimpresso en Berofsky, 1966: 135-159.
- Broncano, Fernando. 2001a. "La educación sentimental. O de la difícil cohabitación de razones y emociones", en *Isegoría* 25: 41-90.
- . 2001b. "Las emociones: territorios intermedios en la mente", en Suplemento 6 de la revista *Contrastes*, Martínez-Freire, Pascual F. (ed) 2001, *Filosofía actual de la mente*: 217-240.
- . Manuscrito. "Acción libre, racionalidad y compatibilismo".
- Calderón de la Barca, Pedro. 1973. *La vida es sueño – El alcalde de Zalamea*, Madrid: EDAF.
- Campbell, C. A. 1951. "Is 'Freewill' a Pseudo-Problem?", en *Mind* LX 240: 446-465. Recogido en parte en Berofsky, 1966: 112-135.
- . 1967. *In Defence of Free Will with Other Philosophical Essays*, London: Allen & Unwin Ltd.
- Chisholm, Roderick M. 1964. "Human Freedom and the Self", en Watson, 1982: 24-35.
- . 1995. "Agents, Causes and Events: The Problem of Free Will", en O'Connor, 1995: 95-100.
- Cicerón, Marco Tulio. 2000. *La adivinación – El bado*, Barcelona: Folio.
- Clarke, Randolph. 1993. "Toward a Credible Agent-Causal Account of Free Will", en O'Connor, 1995: 201-215.
- . 1996. "Agent-Causation and Event-Causation in the Production of Free Action", *Philosophical Topics* 24: 19-48.
- Corbí, Josep E. 2003. *Un lugar para la moral*, Madrid: Antonio Machado.
- Corneille. 1936. *Le Cid*, Paris: Hachette, 1974.
- Cruz, Manuel. 2002. *Filosofía Contemporánea*, Madrid: Taurus.
- Cummins, Robert. 1980 "Culpability and Mental Disorder", *Canadian Journal of Philosophy* X: 207-232.
- D'agostini, Franca. 2000. *Analíticos y continentales*, Madrid: Cátedra.
- Davidson, Donald. 1995. *Ensayos sobre acciones y sucesos*, Barcelona: Crítica.
- . 1963. "Actions, Reasons and Causes", *The Journal of Philosophy* LX, 23: 685-700. Traducido en Davidson, 1995: 17-36.
- . 1992. *Mente, mundo y acción*, Moya, C. ed. Barcelona: Paidós.
- Deleuze, Gilles. 1969. *Logique du sens*, Paris: Les Éditions de Minuit.
- Dennett, Daniel. 1973. "Mechanism and Responsibility", en Honderich, Ted. ed. *Essays on Freedom of Action*. London: Routledge & Kegan Paul. Reimpresso en Watson, 1982: 150-173.
- . 1984. "I Could Not Have Done Otherwise – So What?", *The Journal of Philosophy* LXXXI, 10: 553-567.

- . 2000. *La libertad de acción*, Barcelona: Gedisa.
- . 2004. *La evolución de la libertad*, Barcelona: Paidós.
- Descartes, René. 1996. *Méditations métaphysiques*, Paris: Quadrige/PUF.
- Dick, Philip K. 1992. *¿Sueñan los androides con ovejas eléctricas?*, Barcelona: Edhasa.
- Diógenes Laercio. 2002. *Vida de los más ilustres filósofos griegos*, Barcelona: Folio.
- Double, Richard. 1988. "Libertarianism and Rationality", en O'Connor, 1995: 57-65.
- Duque, Felix. 1998. *Historia de la Filosofía Moderna. La Era de la Crítica*, Madrid: Akal.
- Dworkin, Gerald. 1988. *The Theory and Practice of Autonomy*, New York: Cambridge University Press.
- Epicteto. 1991. *Enquiridión*, Barcelona: Anthropos.
- . 2001. *Disertaciones por Arriano*, Madrid: Gredos.
- Epicuro. 1998. *Obras*, Madrid: Altaya.
- Erasmus de Rotterdam. 1956. *Obras escogidas*, Madrid: Aguilar.
- Estoicos antiguos. 2002. *Fragmentos*, Madrid: Gredos.
- Fischer, J. M. & Ravizza, M. eds. 1993. *Perspectives on Moral Responsibility*, Ithaca: Cornell University Press.
- . 1998. *Responsibility and Control: A Theory of Moral Responsibility*, Cambridge: Cambridge University Press.
- Fischer, John Martin. 1994. *The Metaphysics of Free Will: An Essay on Control*, Cambridge: Blackwell.
- . 1996. "A New Compatibilism", *Philosophical Topics* 24: 49-66.
- . 1999. "Recent Work on Moral Responsibility", *Ethics* 110 (October): 93-139.
- Frankfurt, Harry. 1969. "Alternate Possibilities and Moral Responsibility", *The Journal of Philosophy*, LXVI, 23. Reimpreso en Frankfurt, 1988: 1-10.
- . 1971. "Freedom of the Will and the Concept of a Person", *The Journal of Philosophy*, LXVIII, 1. Reimpreso en Frankfurt, 1988: 11-25.
- . 1987. "Identification and Wholeheartedness", en Schoeman, F. ed. *Responsibility, Character and the Emotions*, Cambridge: Cambridge University Press. Reimpreso en Frankfurt 1988: 159-176.
- . 1988. *The importance of what we care about*, Cambridge: Cambridge University Press.
- Gibbard, Allan. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*, Cambridge: Harvard University Press.
- Gilson, Etienne. 1989. *La filosofía en la Edad Media*, Madrid: Gredos.
- Ginet, Carl. 1989. "Reasons Explanation of Action: An Incompatibilist Account", en O'Connor, 1995: 69-93.
- . 1990. *On Action*, Cambridge: Cambridge University Press.
- . 1996. "In Defense of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Argument Convincing", *Philosophical Perspectives* 10: 403-417.

- . 1997. “Freedom, Responsibility, and Agency”, *Journal of Ethics* 1: 85-98.
- . 2002. “Reasons Explanations of Action: Causalist versus Noncausalist Accounts”, en Kane 2002: 386-405.
- Gómez Caffarena, José. 1970. *Metafísica trascendental*, Madrid: Revista de Occidente.
- Gómez, César. 2003. *Significado y libertad*, Madrid: Siglo XXI de España.
- Habermas, Jürgen. 2002. *Acción comunicativa y razón sin transcendencia*, Barcelona: Paidós.
- Hegel, Georg Wilhelm Friedrich. 1989. *Lecciones sobre la filosofía de la historia universal*, Madrid: Alianza.
- . 1999. *Phénoménologie de l'Esprit* I et II, Paris: Aubier.
- Heidegger, Martin. 1927. *El ser y el tiempo*, Madrid: Fondo de Cultura Económica, 1971.
- Hernández, Javier. 2003. [Publicado en 2004] “Una versión pragmatista del concepto de responsabilidad moral”, *Crítica* Vol. 35, 105 (diciembre): 3-24.
- Hobart, R. E. 1966. “Free Will as Involving Determination and Inconceivable Without It”, en Berofsky 1966: 63-95.
- Hobbes, Thomas. 1966. *The English Works of Thomas Hobbes*, Scientia Aalen: Darmstadt.
- . 1992. *Leviatán*. Vol. I. Madrid: Alianza.
- Hospers, John. 1966. “What Means This Freedom?”, en Berofsky, 1966: 126-142.
- Hughes, G. E. & Creswell, M. J. 1996. *A New Introduction to Modal Logic*, New York: Routledge.
- Hume, David. 2000. *A Treatise of Human Nature*, Oxford: Oxford University Press.
- . 1980. *Investigación sobre el conocimiento humano*, Madrid: Alianza.
- . 1993. *Investigación sobre los principios de la moral*, Madrid: Alianza.
- James, William. 2002. *Pragmatismo*, Barcelona: Folio.
- Kane, Robert. 1985. *Free Will and Values*, Albany: State University of New York Press.
- . 1989. “Two Kinds of Incompatibilism”, *Philosophy and Phenomenological Research*, 50: 219-254. Reimpreso en O'Connor, 1995: 115-150.
- . 1996. *The Significance of Free Will*, Oxford: Oxford University Press.
- . 2001a. ed. *Free Will*, Oxford: Blakwell.
- . 2001b. “Free Will: New Directions for an Ancient Problem”, en Kane 2001a: 222-247.
- . 2002. ed. *The Oxford Handbook of Free Will*, Oxford: Oxford University Press.
- Kant, Immanuel. 1978. *Crítica de la razón pura*, Madrid: Alfaguara.
- . 1989. *Fundamentación de la metafísica de las costumbres*, Madrid: Tecnos.

- . 2000. *Crítica de la razón práctica*, Madrid: Alianza.
- Kierkegaard, Soren. 1973. *Oeuvres complètes*, Vol. VII, Paris: L'Orante.
- Leibniz, G. W. 1959. *Opera Philosophica Omnia (qua exstant latina gallica germanica)*, Aalen: Scientia.
- Locke, John. 1998. *An Essay Concerning Human Understanding*, Kent: Wordsworth Editions.
- Lucas, J. R. 1970. *The Freedom of the Will*, Oxford: Clarendon Press.
- Lucrecio. 1999. *De la naturaleza de las cosas*, Madrid: Cátedra.
- Lutero, Martín. 2001. Teófanés Egido ed. *Obras*, Salamanca: Sígueme.
- Marco Aurelio. 2001. *Meditaciones*, Madrid: Gredos.
- McDowell, John. 1996. *Mind and World*, London: Harvard University Press.
- McKenna, Michael. 1998a. "The Limits of Evil and the Role of Moral Address: A Defense of Strawsonian Compatibilism", en *Journal of Ethics* 2: 123-142.
- . 1998b. "A Speaker-Meaning Theory of Moral Responsibility", ponencia presentada en el *Twentieth World Congress of Philosophy*, Boston (Mass.), August, 1998: 1-9.
- Melden, A. J. 1966. "Free Action", en Berofsky, 1966: 198-220.
- Mele, Alfred R. 1996. "Soft Libertarianism and Frankfurt-Style Scenarios", en *Philosophical Topics* 24: 123-141.
- Melville, Herman. 2004. *Moby Dick I y II*, Madrid: El País.
- Nagel, Thomas. 1979. "Moral Luck", en Nagel, T. *Mortal Questions*, Cambridge: Cambridge University Press. Reimpreso en Watson, 1982: 174-186.
- . 1995. "The Problem of Autonomy". Extracto del cap. 7 de Thomas Nagel, *The View From Nowhere* (New York: Oxford University Press, 1986). En O'Connor, 1995: 33-42.
- Nietzsche, Friedrich. 1997. *La genealogía de la moral*, Madrid: Alianza.
- . 2003. *Más allá del bien y del mal*, Madrid: Edimat.
- Nowell-Smith, P. H. 1948. "Freewill and Moral Responsibility", *Mind* 57: 45-61.
- Nozick, Robert. 1981. "Choice and Indeterminism". Extracto de Robert Nozick, *Philosophical Explanations* (Cambridge: Harvard University Press, 1981). En O'Connor, 1995: 101-114.
- Ocaña García, Marcelino. 1995. *Molina (1535-1600)*, Madrid: Ediciones del Orto.
- O'Connor, Timothy. 1993. "Indeterminism and Free Agency: Three Recent Views", *Philosophy and Phenomenological Research* 53: 499-526.
- . 1995. ed. *Agents, Causes and Events*, Oxford: Oxford University Press.
- . 1996. "Why Agent Causation?", *Philosophical Topics* 24: 143-151.
- Oshana, Marina. 1997. "Ascriptions of Responsibility", *American Philosophical Quarterly*
- Pereboom, Derk. 2000. "Alternative Possibilities and Causal Histories", en Tomberlin, 2000: 119-137.

- . 2001. *Living without Free Will*, Cambridge: Cambridge University Press.
- Pereda, Carlos. 1994. *Vértigos argumentales. Una ética de la disputa*, Barcelona: Anthropos. Especialmente el capítulo II-B. “El dilema fatal”.
- Platón. 2000. *Diálogos III*, Madrid: Gredos.
- . 2000. *Diálogos IV*, Madrid: Gredos.
- Raine, A. & Sanmartín, J. eds. 2000. *Violencia y psicopatía*, Barcelona: Ariel.
- Rawls, John. 1971 [1999]. *A Theory of Justice* (Revised Edition), Cambridge: Harvard University Press. Especialmente el capítulo VIII “The Sense of Justice”.
- Reale, G. & Antiseri, D. 1991. *Historia del pensamiento filosófico y científico I*, Barcelona: Herder.
- . 1992. *Historia del pensamiento filosófico y científico II*, Barcelona: Herder.
- . 1995. *Historia del pensamiento filosófico y científico III*, Barcelona: Herder.
- Reid, Thomas. 1983. Georg Olms Verlag ed. *The Works of Thomas Reid*, Edimburgo: Hamilton.
- Ressler, R. K. & Schachtman, T. 1995. *El que lucha con monstruos*, Barcelona: Planeta.
- Rowe, William L. 1989. “Causing and Being Responsible for What Is Inevitable”, *American Philosophical Quarterly* 26: 153-159.
- Ricoeur, Paul. 1950. *Le Volontaire et l'Involontaire*, Paris : Aubier-Montaigne.
- Russell, Bertrand. 1993. *Ensayos filosóficos*, Madrid: Alianza.
- Russell, Paul. 1992. “Strawson’s Way of Naturalizing Responsibility”, en *Ethics* 102: 287-302.
- Sartre, Jean Paul. 1943. *El ser y la nada*, Madrid: Alianza, 1989.
- . 1946. *L’existencialisme est un humanisme*, Paris: Gallimard, 1996.
- Schelling, Friedrich Wilhelm J. 1989. *Investigaciones filosóficas sobre la esencia de la libertad humana y los objetos con ella relacionados*, Barcelona: Anthropos.
- Schlick, M. 1939. “When Is a Man Responsible”, en 1939, (David Rynin transl.), *Problems of Ethics*, New York: Prentice-Hall. Recogido en Berofsky, 1966: 54-62.
- Schopenhauer, Arthur. 2000. *Sobre la libertad de la voluntad*, Madrid: Alianza.
- . 1994. *Los designios del destino*, Madrid: Tecnos, 1994.
- . 2002. *El mundo como voluntad y representación*, I y II & III y IV, Barcelona: Folio.
- Séneca, Luciano Anneo. 1998. *Diálogos*, Barcelona: Altaya.
- Sófocles. 1988. *Áyax, Las Traquinias, Antígona, Edipo Rey*, Madrid: Alianza.
- Spinoza, Baruch. 1988a. *Ética*, Madrid: Alianza.
- . 1988b. *Correspondencia*, Madrid: Alianza.
- Stern, Lawrence. 1994. “Freedom, Blame, and Moral Community”, *The Journal of Philosophy* 71: 72-84.
- Strawson, Galen. 1986. *Freedom and Belief*. Oxford: Clarendon.
- Strawson, Peter F. 1959. *Individuals: An Essay in Descriptive Metaphysics*, London: Methuen.

- . 1962. "Freedom and Resentment", *Proceedings of the British Academy*, vol. 48: 187-211. Reimpreso en Fischer & Ravizza 1993: 45-66.
- . 1985. *Skepticism and Naturalism: Some Varieties*, New York: Columbia University Press.
- Stuart Mill, John. 1974. *On liberty*, London: Penguin Books.
- Taylor, Charles. 1976. "Responsibility for Self", en Watson, 1982: 111-126.
- Taylor, Christopher & Dennett, Daniel. 2002. "Who's afraid of determinism? Rethinking causes and possibilities", en Kane 2002: 257-277.
- Taylor, Richard. 1982. "Agent and Patient", *Erkenntnis* 18: 111-126.
- Thiebaut, Carlos. 1998. "Responsabilidades cosmopolitas", *La balsa de la medusa* 48: 3-16.
- . 2000a. "Juzgar la acción, rechazar el daño e imputar responsabilidad (sobre la supuesta crisis de la razón moral)", *Taula, quaderns de pensament* 33-34: 15-28.
- . 2000b. "Ponerse en el lugar de(l) otro. (Sobre la atribución de autonomía)", en Bertomeu, Gaeta & Vidiella (comps) 2000: 19-50.
- Tomberlin, James E. ed. 2000. *Action and Freedom*. Oxford: Oxford University Press.
- Urmson, J. O. & Warnock, G. eds. 1961. *Austin: Philosophical Papers*, Oxford: Clarendon Press.
- Valdecantos, Antonio. 2001. "Emociones Responsables", *Isegoría* 25: 63-90.
- Valdés, L. M. ed. 1995. *La búsqueda del significado*, Madrid: Tecnos.
- Van Inwagen, Peter. 1975. "The Incompatibility of Free Will and Determinism", en Kane, 2001: 74-82.
- . 1983. *An Essay on Free Will*. Oxford: Clarendon.
- . 1985. "On Two Arguments for Compatibilism", en *Analysis* 45: 161-163.
- . 1989. "When Is the Will Free?", en O'Connor, 1995: 219-238.
- . 1998. "The Mystery of Metaphysical Freedom", en Van Inwagen, P. & Zimmerman, D. eds. *Metaphysics: The Big Questions*, Oxford: Blackwell. Reimpreso en Kane, 2001: 189-195.
- . 2000. "Free Will Remains a Mystery", *Philosophical Perspectives* 14: 1-19.
- Van Straaten, Zak. ed. 1980. *Philosophical Subjects: Essays presented to P. F. Strawson*. Oxford: Clarendon.
- Voltaire. 1985. *Diccionario filosófico*, Madrid: Akal.
- Wallace, Robert Jay. 1994. *Responsibility and the Moral Sentiments*. Cambridge: Harvard University Press.
- Watson, Gary. 1975. "Free Agency", reimpreso en Watson 1982: 96-110.
- . 1982. *Free Will*. Oxford: Oxford University Press.
- . 1987. "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme", reimpreso en Fischer & Ravizza, 1993: 119-148.
- . 1996. "Two Faces of Responsibility", *Philosophical Topics* 24: 227-248.
- Weatherford, Roy. 1991. *The Implications of Determinism*. London: Routledge.

- White, V. Alan. 1990. "How to Mind One's Ethics: A Reply to Van Inwagen", *Analysis* 50: 33-35.
- Wittgenstein. 1980. *Investigaciones filosóficas*, México: Instituto de Investigaciones Filosóficas de la UNAM.
- . 1988. *Sobre la certeza*, Barcelona: Gedisa.
- . 1997. *Ocasiones filosóficas*, Madrid: Cátedra.
- Wolf, Susan. 1981. "The Importance of Free Will", reimpreso en Fischer & Ravizza, 1993: 101-118.
- . 1987. "Sanity and the Metaphysics of Responsibility", reimpreso en parte en Kane, 2001: 145-163.
- . 1990. *Freedom within Reason*. New York: Oxford University Press.
- Zimmerman, Michael J. 1997. "A Plea for Accuses", *American Philosophical Quarterly* 34: 229-243.